

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2002, 04:55:36 ; Search time 1761.63 Seconds  
(without alignments)  
11791.247 Million cell updates/sec

Title: US-09-732-680a-1  
Perfect score: 1539  
Sequence: 1 attcgacagcagggatcacatc.....cttacaacaaaaaaaaa 1539

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estinu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hlc: \*  
9: gb\_estl: \*  
10: gb\_est2: \*  
11: gb\_hlc: \*  
12: gb\_gss: \*  
13: em\_gss\_hum: \*  
14: em\_gss\_inv: \*  
15: em\_gss\_pln: \*  
16: em\_gss\_vtc: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	400.8	31.9	890	12 CNS013TY	AL103312 Drosophila
2	436.8	28.4	448	9 AT542842	AT542842 SD09326.5
C 3	328	21.3	361	9 AM944588	AM944588 SD09326.5
4	79.6	5.2	891	9 A1110117	A1110117 GH09582.5
5	74	4.8	669	9 BB653265	BB653265 BB653265
6	69.4	4.5	641	9 BB626219	BB626219 BB626219
7	67.6	4.4	554	10 BE664308	BE664308 148713 MA
8	67.6	4.4	755	9 AM914206	AM914206 EST345510
9	66.8	4.3	479	9 AA636687	AA636687 v15f05.r
10	66.6	4.3	837	10 B114221	B114221 602862588
11	65.4	4.2	713	10 BF309251	BF309251 601890196
12	65	4.2	687	10 BB639700	BB639700 BB639700
13	64.8	4.2	654	9 BB626210	BB626210 BB626210
14	64.4	4.2	660	10 BF307986	BF307986 601894231
15	64.4	4.2	806	10 BF305214	BF305214 601892667
16	64.4	4.2	1052	9 AL530299	AL530299 AL530299
17	63.6	4.1	566	9 AA596914	AA596914 v021f03.r

18	63.6	4.1	651	9 BB613689	BB613689 BB613689
19	62.6	4.1	497	9 AA636794	AA636794 v16d12.r
20	62.2	4.0	494	9 AA697326	AA697326 HL02253.5
21	62.2	4.0	754	12 CNS02R2B	AL210044 Tetradon
22	61	4.0	864	9 A0120692	A0120692 AU120692
23	60	3.9	619	10 BM492021	BM492021 p9p2n.pk0
24	58.8	3.8	495	9 AA692456	AA692456 vt21e02.r
25	58	3.8	531	10 BJ065441	BJ065441 BJ065441
26	58	3.8	615	10 BJ032622	BJ032622 BJ032622
27	58	3.8	632	10 BJ035634	BJ035634 BJ035634
28	58	3.8	682	10 BG018521	BG018521 daa46h12.
C 29	57.8	3.8	856	12 CNS04DRH	AL286118 Tetradon
C 30	57.4	3.7	585	12 AZ305495	AZ305495 IM0006N11
C 31	56.6	3.7	443	12 FR0034247	AL030614 Fugu rubr
C 32	56.2	3.7	564	10 BM035632	BM035632 fu18a12.y
C 33	56.2	3.5	694	10 B114994	B114994 602944249
C 34	54.6	3.5	589	9 BB610036	BB610036 BB610036
C 35	54	3.5	438	10 BF544638	BF544638 UI-R-BT0-
C 36	53.4	3.5	428	9 AV633191	AV633191 AV633191
C 37	53.4	3.5	773	10 BF261981	BF261981 HV.CFA000
C 38	53.4	3.5	822	10 BF301720	BF301720 602033423
C 39	53.2	3.5	563	10 B1195946	B1195946 602756234
C 40	52.8	3.4	976	12 BH149983	BH149983 ENTOD93TF
C 41	52.6	3.4	622	10 BE382752	BE382752 601298420
C 42	52.2	3.4	809	10 BG404575	BG404575 602420721
C 43	52	3.4	468	10 BG553005	BG553005 ddb82e09.
C 44	52	3.4	982	12 CNS03F40	AL241233 Tetradon
C 45	51.8	3.4	556	9 AW721149	AW721149 83301H08

## ALIGNMENTS

RESULT 1  
CNS013TY/C  
LOCUS  
DEFINITION CNS013TY 890 bp DNA linear GSS 28-JUL-1999  
Drosophila melanogaster genome survey sequence SP6 end of BAC  
BACN10F18 of DrosBAC library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.

ACCESSION AL103312  
VERSION AL103312.1 GI:5614923

KEYWORDS  
SOURCE

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 890)

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 Evry Cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
- Web : www.genoscope.cns.fr  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome project (EDGP) -  
http://www.edgp.ebi.ac.uk. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billaud at CEPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelobAC11.

FEATURES

Source

Location/Qualifiers

1..890

/organism="Drosophila melanogaster"

/plasmid="pBelobAC11"

/db\_xref="taxon:7227"

/clone\_lib="DrosBAC"

/clone="BACN10F18"

/note="end : SP6"

BASE COUNT 205 a 193 c 240 g 198 t 54 others  
ORIGIN

Query Match 31.9%; Score 490.8; DB 12; Length 890;  
 Best Local Similarity 95.3%; Pred. No. 2.6e-87;  
 Matches 502; Conservative 10; Mismatches 14; Indels 1; Gaps 1;

QY 352 cgtatggagagagagagagcgctgtgtgcaacccgtcacatatatgacaatcagcaatc 411  
 |||||  
 Db 527 CGATGGAGAGCGAGGANCMTTGTCAACCCGTCAAAATATGCAACATCACCGAGATC 468

QY 412 actttagatccagcagagctgtgaccccccaatcacactctcaacgagcagagat 471  
 |||||  
 Db 467 ACTTTGAAGTCCACGAGAGTCTGAGACCCCAATATCACACTTTCAACGGGAGAAAGT 408

QY 472 ggcctgtagcgcgaacccaggtacc-ctcagccagatgagccactccggtgagtc 530  
 |||||  
 Db 407 GGCTGTATGGCCGAACCCAGGTGACCTTCAGCCAGATGGCCACTTCGGTGGATGCC 348

QY 531 tccagccgtgtacacggcctactgtggaactcaatgtctaacaggcccccagcaagca 590  
 |||||  
 Db 347 TCCAGCCGTGTACCGGCTACTGCGAACTCAACTTCTCAACTGGCCCAAGCAAGCA 288

QY 591 gacgtgaagtgaagatcgctcctgagcgctgaagtcgctcctgcgagagagcagc 650  
 |||||  
 Db 287 GAGGTGAAGTGAAGATCGGCTCTCTGGGCTGAAGTCTCTCTGCCGAGAAAGCAGC 228

QY 651 ggcgagagagagagtcctctgacacagagacgtgtcagtcacaggagtggaatcgt 710  
 |||||  
 Db 227 GCGCAGAGAGAGAGTCCCTTACACGACGACCTGTCTAGTACCGGAGTGGAAATCGT 168

QY 711 ggcctgagagagagcactgtgtcagtcagtggaactagcgtcagtggaactgtgac 770  
 |||||  
 Db 167 GGACTCCGAGACCCACTTTPICACTAGAGACTACTAGCGCTACATGGAATGATTTGAC 108

QY 771 ggcctcagcgcgcctcctccatgtacacgagcgtcatctacacacgcgctcgtcatcgt 830  
 |||||  
 Db 107 GGGTACAGGGGGCGTCCCAATGTACAGGGCGGTATGTACACNCCGCTTCTCTTNT 48

QY 831 catcctggcctctcagccttctgtgtcctccacatggcgagc 877  
 |||||  
 Db 47 CTCTCTBSACCTCTTACCTTGTGGCTGCCCTCCACATGKTGCGCG 1

RESULT 2  
 A1542842 448 bp mRNA linear EST 23-APR-2001  
 LOCUS  
 DEFINITION SP09326.3prime SD Drosophila melanogaster Schneider L2 cell culture  
 PORT2 Drosophila melanogaster cDNA clone SP09326 5 similar to  
 G511822: F880011822 'ion channel' located on: 2L 21C5-21C5;:  
 04/13/2001, mRNA sequence.

ACCESSION A1542842  
 VERSION A1542842.2 GI:13771762  
 KEYWORDS EST.  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 448)  
 Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,  
 Lewis, S. and Rubin, G. M.  
 BDGP/HMII Drosophila EST Project  
 .unpublished (2001)  
 On Mar 19, 1999 this sequence version replaced gi:4460215.  
 Other ESTs: SD09326.3prime  
 Contact: Stapleton, M.  
 BDGP  
 Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798  
 Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
 hit genomic AEO03589: arm:2L [301580,604052]  
 estimated-cyto:21B7-21C7: 04/13/2001  
 Plate: SD.93 row: C column: 2  
 High quality sequence stop: 366

FEATURES  
 source POLYA=NO.  
 location/Qualifiers  
 1..448  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /clone="SD09326"  
 /clone\_lib="SD Drosophila melanogaster Schneider L2 cell  
 culture PORT2"  
 /lab\_host="DH5-alpha"  
 /note="Vector: PORT2; Site 1: EcoRI; Site 2: XhoI; Sized  
 fractionated cDNAs were directly ligated into PORT2.  
 Plasmid cDNA library."  
 BASE COUNT 115 a 140 c 116 g 77 t  
 ORIGIN

Query Match 28.4%; Score 436.8; DB 9; Length 448;  
 Best Local Similarity 98.4%; Pred. No. 1.3e-76;  
 Matches 441; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7 cagcaggtatcatccgaaacaaagcgcgctgaaacatgacagactcccaagataag 66  
 |||||  
 Db 1 CAGCACCGTACATCCGAACAAAGGCCGCTGAACATGACAGAGACTCCCAAGATTAAG 60

QY 67 gcaccagttccggtcctgagctgcactactgctgcaatgctaattgagatgctctt 126  
 |||||  
 Db 61 GCACCAATTCGCCGCTCTGAGCTGCACACTGCTGCAAAATGCTAATGGGATGCTTCTT 120

QY 127 atggagtgatcttcgtgccaagcgccactgcccagccgagaccccaagaagccaatgctc 166  
 |||||  
 Db 121 ATGGGCTGTACTTCGCTGCGAGGCGCCACTGCCACCCCGAGACCCCAAGACGCCAATGTC 180

QY 187 aagccctgtagatgcctcccaagcgcgctgttcaacgaactcagacagcagatgtgcaagc 246  
 |||||  
 Db 181 AAGGCGCTGATGATCCCTCCACGCGGCTGTACGAACTACGACAGCATGTGCACCG 240

QY 247 gtgtccaagaaaccccaagaaagtgctccctggaatggtgtaacctacatagacatc 306  
 |||||  
 Db 241 GTGTTCCAAAGAACCCCAAGAAAGTGTCCCTGGAGATGTTGTCACCTACATGACATC 300

QY 307 gacgagttgaacgcaagcgtgacccactgctggtgtaattccgattgagagagag 366  
 |||||  
 Db 301 GACGAGTTGAACGCAAGCTGACCCACTGCTGCTGATCTCCATGAGAGAGAGAG 360

QY 367 gaggcgctgtggaacgcgtcaacatagacaacatcagcagatcttgaatgcagc 426  
 |||||  
 Db 361 GAGGCGCTGTGGCAGACCGTCAATATGACAACTACGACAGATCACTTTGAAGTCCAGC 420

QY 427 gaggctggaaccccccaatcacactct 454  
 |||||  
 Db 421 GAGGTCTGAGACCCCAATCACACTCT 448

RESULT 3  
 AW944588 361 bp mRNA linear EST 05-JUL-2001  
 LOCUS  
 DEFINITION SP09326.3prime SD Drosophila melanogaster Schneider L2 cell culture  
 PORT2 Drosophila melanogaster cDNA clone SP09326 3 similar to  
 AC004573: Drosophila melanogaster, chromosome 2L, region 21C5-21D1,  
 p1 clone DS07610, complete sequence, mRNA sequence.

ACCESSION AW944588  
 VERSION AW944588.1 GI:8122336  
 KEYWORDS EST.  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 361)  
 Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,  
 Lewis, S. and Rubin, G. M.  
 BDGP/HMII Drosophila EST Project  
 unpublished (2001)

COMMENT Other-ESTs: SD09326.5prime

Contact: Stapleton, M.

BDCP

Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST/estefruitfly.berkeley.edu>  
Based upon the presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this clone was polyadenylated. The resulting Poly-T sequence has been removed. hit genomic AE003589: Drosophila melanogaster genomic scaffold 14200013386046 section 15 of 16, complete sequence.: 03/19/2001  
Plate: SD.93 row: C column: 2  
High quality sequence stop: 281.

#### FEATURES

location/Qualifiers  
1..361  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="SD09326"  
/clone\_1lb="SD Drosophila melanogaster Schneider U2 cell culture port2"  
/lab\_host="DH5-alpha"  
/note="Vector: port2; Site\_1: EcoRI; Site\_2: XhoI; Sized fractionated cDNAs were directly ligated into port2. Plasmid cDNA library."  
BASE COUNT 80 a 98 c 102 g 80 t 1 others  
ORIGIN

Query Match 21.3%; Score 328; DB 9; Length 361;

Best Local Similarity 94.2%; Pred. No. 4.7e-55;

Matches 340; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1151 gcgagtcgacgagcgagagaagacaaagatgtagcagcaccgtaagagagcgagc 1210

DB 361 GCGAGTCGACGAGCGGAGAAACCAAGAGATGAGACACACCGGACGAGGCGGACG 302

QY 1211 agcaagagtcagtcgctgggcatcaacacacccagagtgccgagcgcaagcgcaac 1270

DB 301 AGCAGAGATCCAGTCCCTGGGCAATCAACACCCGAGGTGCGGCGGCGCAAGGCCAAC 242

QY 1271 agtcgacgagcgctgtgtgccaacgcccgtgagcgcgcatctctctcaagc 1330

DB 241 AGTTCGATTGGGCGATGTGGCCACCGCGTAGACCGCATTCCTTCCTTCACGCG 182

QY 1331 tggccttccatcttgctgcatcagtgctccgtgtagggatgctcgagactcaagcgca 1390

DB 181 TGGCCTTCCTCATCTGGCCATCAGTGTCCGTATAGGAGTCTCGAGACTCAAGGCGCA 122

QY 1391 catcccaagcagtgcgacacttgacatgttgcatcttgcatcttgcatcttgcatct 1450

DB 121 CATCCCAAGCCAGTGCACACTTTGACACTAGTTTTCATTTTGCATTTTATGATTTAATG 62

QY 1451 tgtgtgcaactataatattatga tgaacctcgatagaataaagacctctg 1510

DB 61 TGTGTGCAACTTATATTCCTTAATGACCANCCCTCGATAGGATTAAGTCTCTGCGC 2

QY 1511 g 1511

DB 1 g 1

#### RESULT 4

LOCUS A110117

DEFINITION 891 bp mRNA linear EST 23-APR-2001  
GH09582.5prime GH Drosophila melanogaster head port2 Drosophila melanogaster cDNA clone GH09582.5 similar to nacr1alpha-96Aa;

FBan0006844 'ion channel' located on: 3R 96A2-96A2.; 04/10/2001, mRNA sequence.

ACCESSION A110117

VERSION A110117.2 GI:13757749

KEYWORDS EST.  
SOURCE fruit fly.

#### ORGANISM

Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 891)

AUTHORS Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G., Lewis, S. and Rubin, G.M.

TITLE BDCP/HMI Drosophila EST Project

JOURNAL Unpublished (2001)

COMMENT On Aug 26, 1998 this sequence version replaced gi:3478441.  
Other-ESTs: GH09582.3prime  
Contact: Stapleton, M.

BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: <http://www.fruitfly.org/EST/estefruitfly.berkeley.edu>  
hit genomic AE003748: arm:3R [20186991,20443598]  
estimated-cyto:96A1-96A20: 04/10/2001  
Plate: GH.95 row: G column: 10  
High quality sequence stop: 647  
POLYA=No.

#### FEATURES

location/Qualifiers  
1..891  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="GH09582"  
/clone\_1lb="GH Drosophila melanogaster head port2"  
/sex="male and female"  
/dev\_stage="adult"  
/lab\_host="DH5 - alpha"  
/note="Organ: head; Vector: port2; Site\_1: EcoRI; Site\_2: XhoI; Sized fractionated cDNAs were directly ligated into port2. Plasmid cDNA library."  
BASE COUNT 209 a 270 c 218 g 194 t  
ORIGIN

Query Match 5.2%; Score 79.6; DB 9; Length 891;

Best Local Similarity 56.7%; Pred. No. 7.2e-06;

Matches 170; Conservative 0; Mismatches 124; Indels 6; Gaps 1;

QY 325 ctgacacccacgctgtgcatctccatgagagagcgcgctggaacg 384

DB 396 CTACACCAACAGCTGTGGCTGAGACAGCAATGCGAGATCAATTAATTAAGTGGATCC 455

QY 385 tcacaatltgaacaatacagcagatcacttgaaagtcacgaggtctggaacccca 444

DB 456 TCGAGATATGGCGCGCTACAGAGCTTATGTGCGCATCCGAGCACATCTGCGCGCAG 515

QY 445 atcaaccttcaagcgagaaagtgtgctgtagccggaaccc-----agtgacc 498

DB 516 ATCGTCTCTACAAACAAATGCGATGCGAGTACGTGTACACCATGACGAAGGCCATC 575

QY 499 ctcaagcaagatgagcatctcggtgagatgcttcagccggtgtaacgacctactgaa 558

DB 576 CTCACATATACCGGCAAGGTGTGAGACTCCGCCGCATCTTAATCCAGCTGTAG 635

QY 559 ctcaacatgctcaactggtcccaagcagcagagctgcaagltgaaatcggtctctg 618

DB 636 ATTGATGTGCGCTACTTCCCTTCATCAGCAGACCTGCTTCATGATGATGCTGCTG 695

#### RESULT 5

LOCUS BB653265

DEFINITION 669 bp mRNA linear EST 26-OCT-2001  
BB653265 RIKEN full-length enriched, adult male hippocampus Mus musculus cDNA clone C630019M18 5', mRNA sequence.

ACCESSION BB653265

VERSION BB653265.1 GI:16487122

KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus







OY	531	tcacacgctgtacacgagcctactggaactcaaatgctcaactgagccccaagaaga	590
Db	451	tccttcgcactcttttaaaagctactggcagactatctgcaccactttcccttgacagaca	510
OY	591	gagctgcaagttgaagatcgctctcg	618
Db	511	gaactgcagcatgaagctggcaccctgg	538
RESULT	8		
LOCUS	AM914206	755 bp	mRNA
DEFINITION	EST345510	Normalized rat brain, Bento Soares	Rattus sp. cDNA clone
VERSION	RG1AB81.5	end, mRNA sequence.	
KEYWORDS	AM914206		
SOURCE	AM914206.1	GI:8079880	
ORGANISM	EST.		
REFERENCE	Rattus sp.		
AUTHORS	Rattus sp.		
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
JOURNAL	1 (bases 1 to 755)		
COMMENT	Lee,N.H., Gloder,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D. Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat Gene Index Unpublished (1998) Contact: Lee, NH The Institute for Genomic Research 9712, Medical Center Drive, Rockville, MD 20850, USA Tel: (301)-838-3529 Fax: (301)-838-0208 Email: nhlee@tigr.org This clone is available through the ATCC, contact the ATCC tel#703-365-2700 for further information Seq primer: M13 Reverse.		
FEATURES	Location/Qualifiers		
SOURCE	1..755		
	/organism="Rattus sp." /db_xref="taxon:10118" /clone="RG1AB81" /clone_id="Normalized rat brain, Bento Soares" /note="Organ: brain; Vector: pT73pac; Site_1: EcorI; Site_2: NotI"		
BASE COUNT	192 a 219 c 176 g 168 t		
ORIGIN			
Query Match	4.4%	Score 67.6:	DB 9: Length 755:
Best Local Similarity	53.1%:	Pred. NO. 0.0017:	
Matches 170:	Conservative 0:	Mismatches 144:	Indels 6: Gaps 1:
OY	305	tcgaagagttgacagcgaagctgaccaccacactgctgctgatactccgatgagagacg	364
Db	84	tgagatgaagtaaacacagatcatgaaacacacactgtgctgaaagcaaatctggatgaact	143
OY	365	aggagcgcgtgtggaacgcgtacaatatgacaacatcacgacagatgcattgaaagcca	424
Db	144	acaaactgtaaaatgaacacccctctgactacacaaagggtgagattgatgcttccgcag	203
OY	425	gagaggtctgagcccccaatcaacatcctttaaagcggaagatgtgctgtatggcgg	484
Db	204	agaaatctggaacacagacatcgtatctgtaacaaacacactgatgggattttccacagttg	263
OY	485	aa-----accaggtgacctcagcgaagatgagcgaactcgcgtggaatgcctcaacgcg	538
Db	264	atgacacagacacaaactctactcaagtacacagagaaactgactttggatccgcggccca	323
OY	539	tgtacacgctactcgcgaactcaacatgctcaactgctgcccacagacaagcagagctgca	598
Db	324	tcttttaagagctcatgcaaaatcgacgtgaacctactgtcccatttgcactccaaaactgca	383

[illegible]

Db 69 GAAACAGCAATGGTGCATATACACTTGAATGCAATCCAGATGACTATGAGAGATGAA 128  
 Qy 405 gcagataccttgaagctcagcgaggtgtgacccccaatcaacttcaac----- 459  
 Db 129 AAAAATTCATCCCTCGGAAAGATCTGGGGCGGAGCGTCTCTATTAACAACGC 188  
 Qy 460 -ggcagcgaaggttgcctgatgtgcccgaaccaggtgacccccaatcaacttcaactt 518  
 Db 189 AGAGCGGACACTTGGCATTTGCAATTCACCAAGGTGCTCTGAGACTACACGGGACAT 248  
 Qy 519 ccggtgatgtcctcagcgcgtgtacacgacctactgtggaactcaactgtcgaactg 578  
 Db 249 CACCTGACACCGCCGACATCTTAAAGCTATGTAGATCATTTGACATCACTTCC 308  
 Qy 579 ccagcagaagaagagctcaagtgtgaatcgtgctcctggggcccgaaagtcgtcctgc 638  
 Db 309 CATCATGACAGACGACATCAGCATGAGCTGGGACCTGAGACTATGACGGCTGTGTGT 368  
 Qy 639 ggaagacgagcagcgag 698  
 Db 369 GGACATTAACCGGAAAGTGAACAGGCG---GACCTGAGTAATCTCATGAGAGCGGGGA 425  
 Qy 699 gtgggaatcgttgactcgcgaagc 723  
 Db 426 GTGGGTGATCAAGAGACCTCGGGCG 450

RESULT 10  
 B114221 837 bp mRNA linear EST 26-JUN-2001

LOCUS 602862588P1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:5021829 5',  
 DEFINITION mRNA sequence.

B114221  
 VERSION B114221  
 KEYWORDS EST.  
 SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 837)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov

Plate: LNCM1840 row: p column: 22  
 High quality sequence stop: 696.

## FEATURES

Source

Location/Qualifiers

1..837

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5021829"

/clone\_id="NIH\_MGC\_17"

/tissue\_type="Rhabdomyosarcoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: muscle; Vector: pORF7; Site\_1: EcoRI;  
 Site\_2: XhoI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using zap-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

## BASE COUNT

ORIGIN

204 a 232 c 213 g 188 t

Query Match 4.3%; Score 66.6; DB 10; Length 837;  
 Best Local Similarity 50.7%; Pred. No. 0.0027;  
 Matches 190; Conservative 0; Mismatches 179; Indels 6; Gaps 1;

Qy 297 catagacatcagcaggttgaacgcggaacgtgacccccaactgtgtcgtgaatctcga 356  
 Db 240 CATCAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 299  
 Qy 357 gagagcgaagagcgcggtgtgcaacgttcaataatgacaacatcagcagatcatt 416  
 Db 300 GGTGATTTACAACTTAAATGGAATGCAATGATGATGATGATGATGATGATGATG 359  
 Qy 417 gaagtcagcagaggtgtgacccccaatcaacttcaactgtgacgcgagagtggtgc 474  
 Db 360 TCCCTTCAACAGATCTGGGCGCCGACCTTGTCTTATTAACAATGAGTGGACTT 419  
 Qy 475 ----ctgattgcccgaaccaggtgacccccaactgtgacccccaactgtgacgcg 530  
 Db 420 TGTATTTGTCAGATTCACCAAGTGTCTCTGAGTACATGAGTACATGAGTACATG 479  
 Qy 531 tccagcgtgtgacacgcgtctactgcgaactcaaatgtcgaactgtgacgcgagagaca 590  
 Db 480 TCCACCATCTTTAAAGCTATGATGATGATGATGATGATGATGATGATGATGATG 539  
 Qy 591 gagctgcaagtgtgaatcgtcctcctggggcctgaaggtgtcgtcgcgagagagagc 650  
 Db 540 GAACGTGACAGATGAAGCTGGGACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 599  
 Qy 651 ggcgagagagagagtc 665  
 Db 600 GGAAGCGACACGAGCC 614

## RESULT 11

BF309251

LOCUS

601890196P1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:4124058 5',

DEFINITION mRNA sequence.

BF309251

VERSION BF309251.1 GI:11256573

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 713)

NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at: Image.llnl.gov

Plate: LNCM1009 row: p column: 23

High quality sequence stop: 711.

## FEATURES

Source

Location/Qualifiers

1..713

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4124058"

/clone\_id="NIH\_MGC\_17"

/tissue\_type="Rhabdomyosarcoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: muscle; Vector: pORF7; Site\_1: EcoRI;  
 Site\_2: XhoI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by



```

location/Qualifiers
1..654
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="g33016116"
/clone_lib="Riken full-length enriched, adult male
diencephalon"
/sex="male"
/tissue_type="diencephalon"
/dev_stage="adult"

```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 660)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)



Fri Jul 19 08:26:18 2002

us-09-732-680a-1.rst

---

Page 11





```

/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
43. .1368
/note="unnamed protein product"
CDS

```

```

/codon_start=1
/protein_id="CAC37851.1"
/db_xref="GI:13939668"
/translation="MTTPKIKAPVSGEPLPLLMINGMLMLGTSVPGATATADPK
NANVALDRIHAGLEFTNPDQVPCPGPNTVSLPMVNTYIDIDELNCKLTHCMIN
LRMRPERWDPQSDYDNTQITLKSEWTPQITTFNDEGLMAETVYTLSDHFR
MPPVAVTAYCELMNLMPHDKOSCKLIGSGKLVVLENTAGEISDHDLVQSP
EMEIVDSRAHFESODYGYMEYTLTAORRSMYTAIVTPASCIYILASMLPMPH
GGEKIMINGLILIVIAAFMLYPAQLLPVLSNTPVLYFVSFLIYLSVSTVEVLY
YLAGKHRRLEPEALRKLHGLHGLTWLISVSTSGEQAETKEMDEPYEADQEO
SSPLGINHTEVPGAKANQFDMLLATADRLSFVSFLILALIRCSV"
BASE COUNT      352 a      476 c      411 g      300 t
ORIGIN

```

```

Query Match      100.0%; Score 1539; DB 6; Length 1539;
Best Local Similarity 100.0%; Pred. No. 2.7e-256;
Matches 1539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

1 attcgacagaggtatcatccgaacaagaagcgctgtgaacaatgaagcgaactcccaag 60
|||||
1 ATTGCGACAGAGGATCATCCGAACAAAGCGCGCTGAACAATGACGAGACTCCCAAG 60
Db
Qy 61 ataaagcaccagttccggtctcgtgactgccaactctgtcgaatgtctaagggatg 120
|||||
Db 61 ATAAAGCACACAGATTCCGGTCCGTGACCTGCACTGCTGCAAAATGCTAATGGGGATG 120
Qy 121 ctctctatggggtgacttcctgctgcaagcgccaactgccaagcggaccaccaagacgccc 180
|||||
Db 121 CTTCCTATGGGGTGTGACTTCCTCGGTGCGACGGCCACTGCGACCGGACCCCAAGAACGCC 180
Qy 181 aatgtcaagacccctgtatcgcctccacgcgcgcgtgttcaagaactacgaacagcgtgtg 240
|||||
Db 181 AATGTCAAGACCCCTGTGATGCGCTCCACGCGCGCTGTTCACGAATACGACAGCATGTG 240
Qy 241 cagcggtgttccaagaagacccccaagcagcgtgtccctggaaatgtgtcaactacata 300
|||||
Db 241 CAGCGGTGTTCGAAGAACCCCGACGAACGTGTCCCTGAAATGTGTGTCACCTACATA 300
Qy 301 gacatcgaagagtgtgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 360
|||||
Db 301 GACATGACGAGATTGAACGGCAAGCTGACCCACTGCTGCTGTAATCTCCGATGGAGA 360
Qy 361 gacgagagcgcgtgtgtgcaaacccgtacaaatgaacaatacgaagcgaagcgaagcgaagcga 420
|||||
Db 361 GACGAGAGCAGCGGTGTGCGCAACCGTCACAATATGACAACTACGAGATCACTTTGAAG 420
Qy 421 tcacgagaggtctggagccccccaatacacatcttcaagcgcggaaggtgtgctgagtg 480
|||||
Db 421 TCACGAGAGGTCTGGAGCCCCCAATACACTCTTCAACGGGAGCAAGGTGGCTGATG 480
Qy 481 gccgaaccgaagtgaacccctcaacccaatgtgacacttcgcgtgtgtatgctccacgcgtg 540
|||||
Db 481 GCCGAACCCAGGTGACCCCTCAACCCAGATGGCACTTCGCGTGAGTGCCTCCAGCGCTG 540
Qy 541 tacagcgactactgtgaactcaacatgtctcaactgtgccccacgaacgaagcgtgcaag 600
|||||
Db 541 TACAGCGCTACTGTGCAACATGCTCAACTGTGCCCCACGAACGAGAGCTGCAAG 600
Qy 601 ttgaagatcgcgtctcgtgggagcgtgaaggtcgtctcgcgaagaagcgaagcgaagaaga 660
|||||
Db 601 TTGAAGATCGGCTCTGGGGGCTGAAGGTCTGCTCCGAGAAACGGACAGGGAAGAAGA 660
Qy 661 gagtcccttgacaacgaacacgtgttcaagtcacgcgaagtgtgaaatctgtgtgctgcga 720
|||||
Db 661 GAGTCCCTTGACACGACGACGCTGTTCACTACCGAGTGGGAATCGTGGACTGCGCA 720
Qy 721 gccacattgtcaagtaagaactactacggtacatagagatatacctgtacggtcagcgg 780
|||||
Db 721 GCCACATTGTCACTAGTACGAGCTACTACGGCTACATGAGTACACTGTGACGGCTACGCG 780
Qy 781 cgctcctcatgtaaacgcgcgtcatctacacccgcgtctcgcgatcgtcatcctgcggcc 840
|||||

```

```

Db 781 CGCTCCCTCATGTACAGCGGCCGTGATCTACACACCCGCGTCTGATGTCATCTGSGCC 840
Qy 841 cctcagcctctgtgctgcctccccaacatggcgcggaagaatcaatgaatcaagcgtg 900
|||||
Db 841 CTTCAAGCCTTGCGGTGCTGCTCCCAATAGGGGGGAGAAATCATATCAAGGCGCG 900
Qy 901 ctcatcatcgtgatcgcgcgtctccatagtacttcgcccagctcctcgaagtgtctc 960
|||||
Db 901 CTCATCATGTGATCCGCCCTTCCTCAAGTACTTGGCCAGCTCCTGCGACGCTGCTGC 960
Qy 961 aacaaatcccaacttgttaactctctacagacccaagcctctgtaacttgagcgtctcc 1020
|||||
Db 961 AACAAATCTCACTGTGTGAATCTTCTACAGACCAAGCTGCTGTATGTGAACGCTTCC 1020
Qy 1021 accatcgtcgaagttcagttctgtacactgtgcacacgaagcacaagaagcgcgtgcg 1080
|||||
Db 1021 ACCATGCTGAGAGTTCTAGTTCGTACGTGCGCACAGGACAGCAAGAGGCGCTGCGCG 1080
Qy 1081 gaagcgtgaagaagcctgtcgaagggacactgtggcagcgtgtgctgtctcgtgtgttc 1140
|||||
Db 1081 GAGGCGCTGAGAAAGCTGTGTCACGGGCACTGGGCACTGGCTGCTGCTGCTGCTGTTC 1140
Qy 1141 agacacactggcgagtcgagcgagcggaagaaagcaaatgagcagcgaagcgtacagag 1200
|||||
Db 1141 AGCACCACTGGCGAGTCTGCGAGCGGAGAAAGACCAAGAGATGAGACGACCCGTACGAG 1200
Qy 1201 gaagcgagcagagcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260
|||||
Db 1201 GAGCGCGAGAGAGAGAGATCTCAGTCCGCTGGGATCAACACACGAGGATGCGGGCGCG 1260
Qy 1261 aagggccaacagctcgaactgtggcgctgtgcacacgcgcgtgtgaecgaattctcctgt 1320
|||||
Db 1261 AAGGCAACACAGTTCGACATGGGGCGCTGTGGCCACCGCGTGAGCCGATTTCTTCGTGT 1320
Qy 1321 tccctgaagcctggccttccatctcgtgacatcagcagcagcagcagcagcagcagcagcagc 1380
|||||
Db 1321 TCCCTGAGCCTGGCCTTCCTCATCTGTGCGCATAGGTGTCTGCTGTGAGGATGCTGAGA 1380
Qy 1381 ctcaagggcacatcccaagaacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1440
|||||
Db 1381 CTCMAAGGCCACATCCCAAGCAAGTGGCACTGTGAACTGATTGTGATTTGGATTTCAAT 1440
Qy 1441 gtaattaatgtgtgtgcaactataatcaatgaatgaagcagcagcagcagcagcagcagcagc 1500
|||||
Db 1441 GATTTAAATGTGTGTGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
Qy 1501 gacctctgcgaatgtctcttcaaaaaaaaaaaaaaaaaa 1539
|||||
Db 1501 GACCTCTGCCGAATGTCTGCTTACAAAAA 1539

```

```

RESULT 2
AY005148
LOCUS      AY005148      1421 bp      mRNA      linear      INV 01-AUG-2001
DEFINITION Drosophila melanogaster nicotinic acetylcholine receptor beta
ACCESSION  AY005148
VERSION     AY005148.1
KEYWORDS   fruit fly, Drosophila melanogaster
ORGANISM   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 1421)
            Grauso, M. and Sattelle, D.B.
            Drosophila melanogaster nicotinic acetylcholine receptor betas3
            subunit (nAChRbeta-21c gene)
TITLE      Unpublished
JOURNAL    2 (bases 1 to 1421)
REFERENCE  2 (bases 1 to 1421)
AUTHORS    Grauso, M. and Sattelle, D.B.
TITLE      Direct Submission
JOURNAL    Submitted (18-JUL-2000) MRC-FGU Human Anatomy and Genetics,

```

FEATURES University of Oxford, South Parks Road, Oxford OX1 3QX, UK  
 source Location/Qualifiers  
 1.1421  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /chromosome="2"  
 /map="2L: 21C5-21C5"  
 /dev\_stage="embryo"  
 45.1370  
 /note="nacr21C; ion channel; transmembrane neurotransmitter"  
 CDS  
 /product="nicotinic acetylcholine receptor beta subunit"  
 /protein\_id="AAF89090.1"  
 /db\_xref="GI:15072331"  
 /translation="MTTPKIKAPVSGPLPILLOMLMGLTSVGTATADPK  
 NANVALDLHAGLFTNYSDVQPVPGOGPTNVSLEMYTYIDIDLNKLTSHGMLN  
 LRMDEERYMOPSOYDNTIOTLKSEWTPQITTFNGEGSLMFTQYTLSDGHER  
 KMPAYVTRAYCELMNLMNPHDKSKLKSGNGLKYLPEENGTARGESLDHDDLVSP  
 EMEIVDSRAHFVSQDYIGMEYTLTAQRSSMTAVIYTPASCIYILASLAWLPDH  
 GEEKIMINGLLIVIAAFMYAQLPVLNNPTPLVIVISLSTLYLSTLYEVLVYL  
 YLATGKHRRLEALRLKLIHGLIGTWLLISVSTGESQAEKTEKDEHPYEADQOE  
 SSPLGINHFEVGAKANQDFMALATAVDRIISFVSFLAFLILAIRCSV"  
 BASE COUNT 311 a 463 c 389 g 258 t  
 ORIGIN  
 Query Match 90.4%; Score 1390.8; DB 3; Length 1421;  
 Best Local Similarity 99.4%; Pred. No. 1.1e-230;  
 Matches 1406; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Db 609 ATCGGCTCTGGGGCCCTGAAGGTCGTCCTGCCGAGAACGGCACGGCGAGAGAGACTCC 668  
 Qy 667 ctgacacgaacgaacacgtgttcagtcacccgaatgggaaatcgtgacatcgagacccac 726  
 Db 669 CTGACACGACGACGACCTGGTTCAGTCAACGGAGTGGGAAATCTGTGACTGCCAGCCAC 728  
 Qy 727 ttgtcagtcaggaactactacacgtctacataggagtaacatctgaacgtctcagcgctcc 786  
 Db 729 TTGTGACGTACGAGTACTACGCTACATGAGTACACTGTACGGGTACGGCGGCTCC 788  
 Qy 787 ttcagttaacagcgctgaatctactacacacccgctcctgtgaatcgtatcccgccctca 846  
 Db 789 TCCATGTACAGGCGCCGTCATCTACACACCCGCGTCTCGCATGCTCATCTGGCCCTCTCA 848  
 Qy 847 gctctcgtcgtcctcccccacatggcgaggaagaatcagtaacacggcgcgtccatc 906  
 Db 849 GCCTTCGGCTGCTCCGCCACATAGGGGGGAGAAATCATATCAAGGGCGTCTATC 908  
 Qy 907 atcgtgatcgcgcttcctcatgtactctgcacagctcctgcgaatgctgtccaacat 966  
 Db 909 ATCGTATCGCCGCTCTCCATGATGACTTTCGCCAGCTCCGACAGTGTCTCAACAAAT 968  
 Qy 967 acccaactgttgtaatctctctacagacacagcctgctgatatctgagcgtctcacatc 1026  
 Db 969 ACTCCACTTGTGGTATCTTCTACAGCACCGCCTGCTGATCTGACGCTCCACCATC 1028  
 Qy 1027 gtccaggtctcagttctcgtacccctggtccacaggaacacacagagcgctccggagagcg 1086  
 Db 1029 GTGAGGTTCTAGTCTGTGTAAGTGTGACGACAGGACAGGACAAAGGGCGCTGCCGAGGCG 1088  
 Qy 1087 ctgagaagctgtctgacagcgacacgtggcgagctgtgctgtctcgtgtgttcacgac 1146  
 Db 1089 CTGAGAAAGCTGTGTGACAGGGGACCTGGGACCTGGCTGCTGCTCGGTTCAGACAC 1148  
 Qy 1147 actgagcagtcgcagagcgaggaagaacaaagagatgagacagacccgtacaggaagcg 1206  
 Db 1149 ACCGGCGAGTGCAGAGCGGAGGAAAGACCAAGAGATGAGACGACCCGTAAGAGAGGCG 1208  
 Qy 1207 gacgagcaggaatccagttccgtctggagcatcaacacacacaggtgcccggcgacaaagcc 1266  
 Db 1209 GACGACGAGGAGTGCAGTCCGCTGGGATCAACACACCGAGGTGGCGGCGCAAGGCG 1268  
 Qy 1267 aaccagttgacatggggcgctggtgcacacgcggtgagacccgattctctgttctctc 1326  
 Db 1269 AACCAAGTTCAGTGGGGGCTGTGGCCACCGCTGGACCGCATTTCTCTGTTCTCTTC 1328  
 Qy 1327 agcctgacctctctcattctcgtgcacatcaggtgtctcgtgtaggaatgctcgaactcaag 1386  
 Db 1329 AGCCTGGCTTCTCTCATTTGCGCATCAGGTGCTCGCTGAGGATGCTCGAGACTCAA- 1387  
 Qy 1387 gccacatcccaagccagctgcgaactctgaactag 1420  
 Db 1388 GCCACATCCCAAGCCAGTGGCAGCTCGAAGCTAG 1421  
 RESULT 3  
 LOCUS AY061622 1646 bp mRNA linear INV 08-NOV-2001  
 DEFINITION Drosophila melanogaster SD09326 full length cDNA.  
 ACCESSION AY061622  
 VERSION AY061622.1 GI:16769901  
 KEYWORDS FLI CDNA.  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 Muscomorpha; Ephyridiidae; Drosophilidae; Drosophila.  
 1 (bases 1 to 1646)  
 Stapleton,M., Brokstein,P., Hong,L., Agbayan,A., Carlson,J.,  
 Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R.,  
 Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,  
 Nuno,J., Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K.,  
 Yu,C., Lewis,S.E., Rubin,G.M. and Celisner,S.



QY 1266 caaccagtcgacgctggcgtctgcccacccgctgaccgcatcttcctgcttcc 1325  
|||||  
Db 1381 CAACAGTGTGACCTGGGGGCTGCTGGCCACGCGGTAGACGGCATTTCTTCTT 1440  
QY 1326 cagcctggtccttcctcattctgcccattcagtgctcgctgtaaggaatgctcgaactca 1385  
|||||  
Db 1441 CAGCTGGCTCTTCATCCCTGGCCATCAGGTGTTCCGTATAGGATCTCGAGACTCAA 1500  
QY 1386 ggcacatcccaacccagtgccactcgaactagtttgatttgctgatttcatt 1445  
|||||  
Db 1501 GGGCAGATCCCAAGCCATGTGGCACTGAACTAGTTTGGATTGGCATTTGATGATT 1560  
QY 1446 taactgtgtgcgaactataatttaagatgaagaccctgataagaaagacct 1505  
|||||  
Db 1561 TAACTGTGTGCTGCAACTATATATTTAATGATGAGACCTCGTATGGAATAAGGACT 1620  
QY 1506 ctgccgaatgtctgcttacaacaaaa 1531  
|||||  
Db 1621 CTGCCGAAAAAATAAAAAA 1646

RESULT 4  
DME318761  
LOCUS DME318761 1334 bp mRNA linear INV 02-AUG-2001  
DEFINITION Drosophila melanogaster mRNA for nicotinic acetylcholine receptor  
beta 3 (Dbeta3) subunit (nAcRbeta-21C gene).  
ACCESSION AJ318761  
VERSION AJ318761.1 GI:15130913  
KEYWORDS nAcRbeta-21C gene; nicotinic acetylcholine receptor beta 3.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscophora; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 1334)  
Lansdel,S.J. and Millar,N.S.  
Unpublished  
2 (bases 1 to 1334)  
Millar,N.S.  
Direct Submission  
Submitted (31-JUL-2001) Millar N.S., Department of Pharmacology,  
University College London, Gower Street, London, WC1E 6BT, UNITED  
KINGDOM

FEATURES  
source Location/Qualifiers  
1..1334  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
5..1330  
/gene="nAcRbeta-21C"  
5..1330  
/gene="nAcRbeta-21C"  
/function="nicotinic acetylcholine receptor subunit"  
/codon\_start=1  
/product="nicotinic acetylcholine receptor beta 3 (Dbeta3)  
subunit"  
/protein\_id="CAC48166.1"  
/db\_xref="GI:15130914"  
translation="MTTTRKIKAPVSGPGLPLIOMLGMILGSLTSPGATATADPK  
NANVKALDRLHAGLFTNYDSQVDFEGTPTNVSLEAVVYIDIDELNGKLTTHCWLN  
LRWDERVWQPSQYDNITITKSSSEVMPQITLFGNDEGLAAEQVLSHDGHR  
LMPPAVYTAACELMLNMPHDKSKLKISGSLKAVLPENGTARGSLDHDLDVQSP  
EMELVDSRAHFVSODYGYMEXYTLTAORSSMYTAVYTPASCTVIALASAFMLPPLM  
GGEKIMINGLIIVIAFLMYFAQDLPEVLSNNTPLVIFYSTSLIYSYIVAYVLY  
YLATGKHRRRLPEALRKLHGLSTWLLSVSTGSOAEKTKEMDHPYEEADEOE  
SSPLGINHTVEPKAKNQFDWALLATVADRISFVSFLAFLIALIRCSV"

BASE COUNT 285 a 433 c 368 g 248 t  
ORIGIN

Query Match 86.4%; Score 1329.2; DB 3; Length 1334;  
Best Local Similarity 99.8%; Pred. No. 4,7e-220;  
Matches 1331; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 39 aacaatgacgacgactcccaagataaagcaaccagttccggtctctgactgccaact 98  
|||||  
Db 1 AACATATGACGACGACTCCCAAGATTAAGGACACACTTTCGGTCTCTGACTGCCACTACT 60  
QY 99 gctgcaaatgcttaatgggagtgcttcttaagggtctgactcgtgccaagcgccaatgctc 158  
|||||  
Db 61 GCTCAATATGCTAATGGGATGCTTCTTATGGGGCTGACTTCCGTCGACGCGCACACTGC 120  
QY 159 caccgcggaaccccaagaagcccaatgltcaaggccctgtaatgctctccagcgccgctgct 218  
|||||  
Db 121 CACGCGGACCCCAAGACGGCAATGTCAAGGCCCTGTGATGCTTCCACGCGGCTGT 180  
QY 219 caagaaatgacgacgactgctgacgctggtgttccaaagaaaccccaagaaagtgctcct 278  
|||||  
Db 181 CACGACACTACGACGACGATGTGCAAGCCGCTTCCAAAGAAACCCCGACGAAGTGTCCCT 240  
QY 279 ggaatggtgtgctacctaataagacatcgaagltgaacggaagctgaaccccaatg 338  
|||||  
Db 241 GGAATAGTGTGTCACACTACATAGACATGACGAGTGTGAACGGCAAGTGAACCCACTG 300  
QY 339 ctgactgaatctccgaatggaagagagagcggtgtgtgcaaccgttcaaatatgaca 398  
|||||  
Db 301 CTGCTGATCTCCGATGAGAGAGACGAGAGCGGTGTGGCAACCGTCAATATGACAA 360  
QY 399 catcaagcagaatcaacttgaagttcgaagcgaaggtctggaacccccaatacacacttcaa 458  
|||||  
Db 361 CATCACGAGATCACTTTGAAGTCCAGCGAGGCTGTGACCCCCCAATACACTCTTCAA 420  
QY 459 cggcgcaagaaagtgtgctgactgacgcaaaacccaaagtgaacccctcagcccaatg 518  
|||||  
Db 421 CGGCGACGAAGTGTGCTGATGGCGCGAAGCCAGGTGACCCCTCAAGCCAGATGGCACTT 480  
QY 519 ccggtggaatgctcctcagcgcgtgtacagcgctactgtggaactcaaatgctcaactggc 578  
|||||  
Db 481 CCGGTGATGCTCTCAGCGGTGTACAGCGCTACTGCAACTCAACATGCTCAACTGTGCC 540  
QY 579 ccacgaacagcagaagctcaagltgaagatcgctcctggggcctgaaagtcgtccctgcc 638  
|||||  
Db 541 CCACGACAGCAGACGACTGCAAGTTGAGATGGCTCTCTGGGGCCTGAAAGTGTCTGCC 600  
QY 639 ggaagaacggaacggaagagagaggtcccttgacacgaacgactggttcagtaacgga 698  
|||||  
Db 601 GGAGAACGCGACGCGAGAGAGAGATGCTTGTACACGACGACGCTGTTCACTACCGCA 660  
QY 699 gtgggaatgctgtgactcgtgagcccaacttgtgaagtaaggaactactacggtctacatg 758  
|||||  
Db 661 GTGGGAATCGTGACCTGTGCGAGCCCACTTGTGATGAGGACATGACGCTACATGGA 720  
QY 759 gtacacttgacgactcgaagcgctcctccatgttacaacgacgctacatacacaccgc 818  
|||||  
Db 721 GTACACTGTGACGGCTGACGGCGCTCTCCATGTACAGGCGCTGATCTACACACCGCG 780  
QY 819 gtctgtacatgctactcctgtgaccttcagcctctgtgctcctcccaatgagcgagga 878  
|||||  
Db 781 GTCTGTGATCGTCAATCCTGGCCCTTCTAGCCTTGTGCTGCTGCCCTCCACATGGGGGGA 840  
QY 879 gaagatcatgttcaacgagcgtctcatcatcatgtgtagtcgcgcctctccatgctactgc 938  
|||||  
Db 841 GAAGATCATGTATCAACGGCTGTCTATATCTGTATCTCCCTCTTCATGATCTTGC 900  
QY 939 ccaagctctgcaatgtgctgtccaaacatctcccaacttgtgtaactcttcaacaacag 998  
|||||  
Db 901 CCAAGCTCTGCGCATGCTGCTCCAAATATCTCCACTTGTGATCTTCTACAGCACAG 960  
QY 999 cctgtgtatcttgagcgctctccacatctgtagagtttctgtaactctgtgcaacag 1058  
|||||  
Db 961 CCTGCTGATCTGAGCGCTCCTCACCATGCTGAGGTTCTAGTTCGTGCTGCGCACAGG 1020  
QY 1059 caagcacaagaagcgctgctgcaagcgctgtagaagagctgtgcaacggaacactgaggac 1118  
|||||  
Db 1021 CAAGCAAGAAGGCGCTGCGGAGGCGCTGAGAAAGCTGTTGACGAGGCACTGGGGAC 1080  
QY 1119 gtgctgtgtctctggtgttcagcaccactggtgagtcgagcgaggaagcaacaaga 1178

Db	1081	gmgcgtgctgctctgcgggtttcaggcaccacggcgagttgcagggggagaaacacacagga	1140
Qy	1179	gatgagcagaagcaccctgtatcagaggaagcggagcagaagatccagttcgcctlyggcatcaa	1238
Db	1141	gatggacacagacacccctgacgagagagggacgacagagagatccacgtccgcctgggcatcna	1200
Qy	1239	ccacacccgaagtgcgcgggcgcgccaagggccaaacagttctgaactlyggcgtctlyggcacgcg	1298
Db	1201	ccacacccgaggggcgcggggcgccaaaggccaaacagttcgactgggggctgctggccaccgc	1260
Qy	1299	catgagacgcgattctctcgtttctcttcaagctcgtgccttcctcatcttgggcataagtg	1358
Db	1261	cctggacccgcgatttccttctcgtttcctttcagccttgggcctttccatttttggccatcaggtg	1320
Qy	1359	ctcgcgtatgaggat	1372
Db	1321	ctccgctgtagggat	1334

RESULT	5	
AC004573		
LOCUS	AC004573	85095 bp DNA
DEFINITION	Drosophila melanogaster, chromosome 2L, region 21C5-21D1, P1 clone DS07610, complete sequence.	
ACCESSION	AC004573	
VERSION	AC004573.1	GI:4204255
KEYWORDS	HTG.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster	

## REFERENCE AUTHORS

TITLE	Sequencing of Drosophila chromosome 2L, region 21C5-21D1
JOURNAL	Unpublished (1998)
REFERENCE	2 (bases 1 to 85095)
AUTHORS	Celniker,S.E., George,R.A., Galle,R., Svitskas,R.R., Hoskins,R.A.

TITLE	Direct Submission
JOURNAL	Submitted (17-APR-1998) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT	On Jan 30, 1999 this sequence version replaced gi:4164118.

## FEATURES

### source

```

/organism="Diosaphina meranogasteri"
/strain="y2; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2L"

```

```

/imap="21c5-21d1"
/clone="P1_DS07610_(D275)"
/clone_lib="P1 library, partial Sau3A in pMS582tec144d10"
BASE COUNT      22289 a 19671 c 20319 g 22816 t
ORIGIN

```

Query Match	67.8%;	Score 1044.2;	DB 3;	Length 85095;
Best Local Similarity	77.2%;	Pred. No. 5.8e-171;		
Matches 1518;	Conservative	0;	Mismatches 13;	Indels 436; Gaps 3

[illegible]

Db 5646 AGTCGCAAGTTGAAGATGCGCTCCTGGGGCCGTGAAGGTGCTGTCGCCGAGACGGCAGC 5705  
 Qy 652 gcgagagagagagtcctctgaccagacagacgtgttcagtcaccggatgtggaatcgtg 711  
 Db 5706 GCGAGAGAGAGTCCCTTGGACACGACGACCTGGTTCCAGTACCGGAGTGGGAATGCTG 5765  
 Qy 712 gactcgagagccacttctgctcagtcagtaactacagctacatgaggtacactctgacg 771  
 Db 5766 GACTCGGAGAGCCACTTGTGTCAGTCAGAGACTACAGGCTCATGAGATGAGTACACTGTGAC 5825  
 Qy 772 gctcagagagagtcctcctcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 831  
 Db 5826 GCTCAGAGGAGGCTCTCTCATGTCATACAGGCGCTCATATCACACCGCGCTCTGATCGTC 5885  
 Qy 832 atcttgccctctcagctctctgctcctcctcctcctcctcctcctcctcctcctcctcctc 891  
 Db 5886 ATCTTGCCCTCTCAGCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5945  
 Qy 892 aacgagctctcactcactcactcactcactcactcactcactcactcactcactcactcactc 951  
 Db 5946 AAGCGCCTGCTCATCATGTCGTCGCGCCTTCTCATGTCGTCGTCGTCGTCGTCGTCGTCG 6005  
 Qy 952 gtgtgttccaaactactcactcgtg----- 978  
 Db 6006 GTGCTGTCCAACTACTCCACTTGTGGTGAGTGTCTCCCATTTGCTCATTCAGTTC 6065  
 Qy 979 ----- 978  
 Db 6066 AGTTCGAAACTGATGCTAAGTACCTTAACATCCGAGANTTTACCGACTTTACCGA 6125  
 Qy 979 -----gtaactctcactcagcagcctctgtaact 1010  
 Db 6126 CCACTCACCGGCGATTCCTATATTTGCAATATTTCTTACAGCACACGCGCTGCTGATCT 6185  
 Qy 1011 gacgctcctcactcactcactcactcactcactcactcactcactcactcactcactcactc 1070  
 Db 6186 GAGGCTCTCCACATCTCGAGGTTCTAGTCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTG 6245  
 Qy 1071 ggcgctcctcagagagcgtctgaaagaagctgctcagtcagtcagtcagtcagtcagtcagtc 1130  
 Db 6246 GCGCTCTCGGAGGCGCTGAAAGAGCTGTTGACAGGAGACCTGGGACAGTGCTGCTGCTG 6305  
 Qy 1131 ctctgttctcagcactcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1190  
 Db 6306 CTGCTGTTCACACACACCGCGAGTGTGCAAGGAGAGAGATGAGAGAGCA 6365  
 Qy 1191 ccgctcagagagagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1250  
 Db 6366 CCCCTAGAGAGGCGGAGAGAGAGTCCAGTCCGCTGGGCAATCAACACACCGAGGT 6425  
 Qy 1251 ggcgagagagagagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 1310  
 Db 6426 GCCGGGCGCCCAAGGCGACAGTTCGACTGAGGCGCTGCTGCGCACCGCGTGGACCGAT 6485  
 Qy 1311 ttcctcgttctcactcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1370  
 Db 6486 TTCCTGCTTCTTCAGCTGCGCTTCCATTCCTGACCATGAGTCTCGTGTAGG 6545  
 Qy 1371 atgtctcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1430  
 Db 6546 ATGTCTGAGACTCAAGGCGACATCCCAAGCCAGTGGCGACCTCAACACTGTTTGGCAAT 6605  
 Qy 1431 ggcgttctcactcactcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1490  
 Db 6606 GCGATTCATATATTTAATGTGTGTGCGAATTTAATATTTAATGTGTGAGACCTCGTA 6665  
 Qy 1491 tgaataaagagactctgctcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1537  
 Db 6666 TGAATAAAGSACCTCTGCGAATGTCTGCTTACAACTAGAGAGGAA 6712

RESULT 6  
 AC017491/c

LOCUS AC017491 142257 bp DNA linear HTG 09-DEC-1999  
 DEFINITION Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*, In ordered  
 pieces.  
 ACCESSION AC017491  
 VERSION AC017491.1 GI:6553495  
 KEYWORDS HTG; HTGS\_PHASE2.  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 142257)  
 AUTHORS Adams, M. and Venter, J.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD, USA  
 COMMENT This sequence was identified as CDL10210988 by the submitter.  
 For more information on this record e-mail to fly@celera.com.  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 FEATURES  
 source  
 1. 142257  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 BASE COUNT 38086 a 33712 c 32464 g 37995 t  
 ORIGIN  
 Query Match 67.8%; Score 1044.2; DB 2; Length 142257;  
 Best Local Similarity 77.2%; Pred. No. 5.5e-171;  
 Matches 1518; Conservative 0; Mismatches 13; Indels 436; Gaps 3;  
 Qy 7 cagcaggtacatccgaagaagcgcgcgtgaacatgacgactcccaagataag 66  
 Db 122926 CAGCAGCATCATCGAAGCAAGGCGCGTGAGCAATGACGAGACTCCCAAGATAAG 122867  
 Qy 67 gcacagttccggttcctcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 126  
 Db 122866 GCACGAGTTCTCGGTCCTCGGAGCTCCACTGCTGCAAAATGTAATGGGATCTTCT 122807  
 Qy 127 atgggagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 156  
 Db 122806 ATGGGCGTGAATTCCTGTCGCGGCGCCACTGTAAAGTCTTACCGGCTCCGAGTTGCATA 122747  
 Qy 157 -----gcacgcgagaccccaagaagccaa 182  
 Db 122746 AGCGCTGATTAAGTGTATCTGAAATCCCTCCAGCCACCGCGAGCCCAAGAGGCCAA 122687  
 Qy 183 tgttaagccttgatgctcccaagcgcgcgtgttcaagactaagcagatgtgca 242  
 Db 122686 TGTAAAGCCTTGATGCTCCACAGCGCGCTGTTCACGAACTACGACAGCATGTGCA 122627  
 Qy 243 ggcggtgtcccaagaagcccccacagacgtgtccctggaatggtgtgcaactacataga 302  
 Db 122626 GCGGCTGTCCAAAGAACCCCAAGAACAGTGTCCCTGGAATGTGTGTACACTACATAGA 122567  
 Qy 303 catcagcagttgaaagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 351  
 Db 122566 CATCGACAGTTGAAGCGCAAGCTGACACCACTGCTGTGATCTCGTAATCTCCAGTTCAGT 122507  
 Qy 352 ----- 351  
 Db 122506 CCTCTCTTCT 122447  
 Qy 352 ----- 351  
 Db 122446 CCTCAGTTTCACTGATGATGAAGCAATTTTGAAGAGCGTTTATCTGAACTGAACTGAA 122387  
 Qy 352 ----- 351  
 Db 122386 GTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 122327

[illegible]

QY	1251	gcggcgcccaaggccaaacgafttcgaagggggqgctgctgagccacgcgctgagaccgat	1310
Db	121246	gccggcgcccaaggccaaacgafttcgaagggggqgctgctgagccacgcgctgagaccgat	121187
QY	1311	ttccctcgcttctcctcagccctgagccttcctcctcattcctgycacatcagtgctccgltgaa	1370
Db	121186	ttccctcgcttctcctcagccctgagccttcctcctcattcctgycacatcagtgctccgltgaa	121127
QY	1371	atgctgaagactcaaggccacatcccaagaagctgagcacttgagactgaatttcacatt	1430
Db	121126	atgctgaagactcaaggccacatcccaagaagctgagcacttgagactgaatttcacatt	121067
QY	1431	gcgaattcaatgatttaattgctgctgagcaactataattatttaagabgaacctgta	1490
Db	121066	gcgaattcaatgatttaattgctgctgagcaactataattatttaagabgaacctgta	121007
QY	1491	tggagataagagacctgcgcgaatgctcgtcttcaaaaaaaaaaaaaa	1537
Db	121006	tggagataagagacctgcgcgaatgctcgtcttcaaaaaaaaaaaaaa	120960

RESULT	7
AC008002/c	
LOCUS	
DEFINITION	AC008002 182726 bp DNA linear INV 06-APR-2001 Drosophila melanogaster, chromosome 2L, region 21P-21E, BAC clone BACR48t08, complete sequence.
ACCESSION	AC008002
VERSION	AC008002.3 GI:13559545
KEYWORDS	HTG.
SOURCE	fruit fly.

ORGANISM  
*Drosophila melanogaster*  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; *Drosophila*.  
1 (bases 1 to 182726)  
Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,  
REFERENCE  
AUTHORS

Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y., An, H., Baldwin, D., Bazzon, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Chame, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, Y., Doup, L.E., Doyle, C., Dresnek, D., Fartan, D., Ferreira, S., Frishe, E., Galle, R.F., Garay, N.S., George, R.A., Gonzalez, M., Honick, J., Hoshins, R.A., Hostin, D., Howland, T.J., Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattei, B., Mosheif, A., McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, R.A., Nunoo, J., Paclol, J., Paradis, V., Park, S., Patel, S., Pfeiffer, B., Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F., Stapleton, M., Strong, R., Svirskas, R., Vector, C., Williams, S.M., Zaveril, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C. Sequencing of *Drosophila* chromosome 2L, region 21D-21E

REFERENCE  
AUTHORS  
2 (bases 1 to 182726)  
Celniker,S.E., Agbavani,A., Arcaina,T.T., Baxter,E., Blazek,R.G.,  
Butenhoff,C., Champé,M., Chavez,C., Chew,M., Ciesiolka,L.,  
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,  
Hoskins,R.A., Houston,K.A., Hummstedt,S.R., Karra,K., Kearney,L.,  
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Madad,P.,  
Moshrefi,A.R., Moshrefi,M., Nixon,K., Paclob,J.M., Park,S.,  
Peiffer,B., Poon,L., Sequeira,A., Sethi,H., Shih,E.,  
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and  
Rubin,G.M.

TITLE  
JOURNAL  
COMMENT  
Direct Submission  
Submitted (09-JUL-1999) Drosophila Genome Center, Lawrence Berkeley  
Laboratory, MS 64-121, Berkeley, CA 94720, USA  
On Apr 6, 2001 this sequence version replaced gl:6532012.

Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory, MS 64-121  
Berkeley, CA 94720  
This sequence was assembled using end sequences from a whole genome  
shotgun and from subclones of this BAC and its neighboring clones.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our sequence  
archive Web site (<http://www.fruitfly.org/sequence/>) or send email



to bdp@fruitfly.berkeley.edu.  
FEATURES  
Location/Qualifiers  
source  
1. .182726

/organism="Drosophila melanogaster"  
/strain="y; cn bw sp"  
/db\_xref="taxon:7227"  
/chromosome="2L"  
/map="21D-21E"  
/clone="BACR48E08 (D843)"  
/clone\_lib="RPCT-98 (Roswell Park Cancer Institute  
Drosophila melanogaster BAC library, partial EcoRI in  
PBac3.6)"  
BASE COUNT 49900 a 42515 c 41297 g 49014 t  
ORIGIN

Query Match 67.8%; Score 1044.2; DB 3; Length 182726;  
Best Local Similarity 77.2%; Pred. No. 5,4e-171;  
Matches 1518; Conservative 0; Mismatches 13; Indels 436; Gaps 3;

7 cagcaggtatcatccgaacaaaggcgctgtgaacatgtacgacgtcccaagataag 66  
Db 112151 CAGCACCCTGATCCGAACAAAGCGCGCTGAACATGACGACACTCCCAAGATAAG 112092  
Qy 67 gcacaaattccggtcctgtgactgcaactactgctgaatgttaactggtggtctct 126  
Db 112091 GCACCAATTCGCGTCTGTGACCTGCCACTGCTGCAATGCTAATGGGATCTCTT 112032  
Qy 127 atggagtgacttcgtgtcgaagcgccact----- 156  
Db 112031 ATGGGCTGACTTCCGTCGACGGCGCAGCTGTAACTGTATACGCTCCGAGTTGCATA 111972  
Qy 157 ----- gcaacgcgcgaaccccaagaagcgcaa 182  
Db 111971 AGCGCGTAAATGATGTATCTGAATCCCTCCACGCGGAGCCCAAGACGCGCAA 111912  
Qy 183 tgcataagcctgcatcgctccacgcgcgcgtgttcacgaactacagacgcgtgtcga 242  
Db 111911 TGTCAAGCGCTGATCGCTCCACGCGGCGCTTTCACGAACATGACACAGCATGTCGA 111852  
Qy 243 gccgggtgtccaaagaaaccccaacgaacgtgtccctgtgaatgtgtgttcacatacaga 302  
Db 111851 GCCGGTGTTCACAAAGAACCCCGACGAACGTCCTGGAATGGTGTGCTACATCATAAGA 111792  
Qy 303 catcgagaattgaacggaagcgtgaccacccactgtgtgtgaatctc----- 351  
Db 111791 CATCGAGAGTGAACGGCAAGCTGACCCACTGCTGTGATCTGTAATCTGTAATCTCAGT 111732  
Qy 352 ----- 351  
Db 111731 CTTCTCTTGTGTCCTCTTCCTTCGTCGAGTACGTGAGAAACGATTAATTCGAGATTTC 111672  
Qy 352 ----- 351  
Db 111671 CTTCAAGTTTCAACACTCTGATTAAGCAATTTTGAAACGTCGTATCTGAACGTAACTGAA 111612  
Qy 352 ----- 351  
Db 111611 GTGTTGTTTTGTTTTTTTTTTTTTTTTTTTGTGAGAGAGTACTATTAAGCTAATAAGTCTA 111552  
Qy 352 ----- 351  
Db 111551 GTGCAATTATGAATATCTGAATATCCAAATCACTAACACAGCAATCGTGTTTCTAG 111492  
Qy 352 cgatggagagacgaagagcggtgtgtgcaacgctcacataatgacacatcacgcagatc 411  
Db 111491 CGATGGAGACAGCAGGACCGCTGTGCAACCGTCACAAATATGACACATCACACAGATC 111432  
Qy 412 acattgaagtcgaagcggtgtgtgaaccccaatcactcttcaacgcggaagaggt 471  
Db 111431 ACTTTGAAGTCCAGCGAGTCTGGACCCCAATCACTCTTCAACGGGAGCAAGAGT 111372  
Qy 472 ggcctgtatgctcgaaacccagtgaccctcagccacgatgscacttcggtgatgct 531

Db 111371 GGCCTGATGGCGCAAAACCCAGGTGACCCCTGACGACATGGCCACTTCGCGTGAATCCCT 111312  
Qy 532 ccagcgtgtacacagcgcctactcgtgaactcaacatgtcactggtcccaacagacag 591  
Db 111311 CCACCCCTGTGACAGCGGCTCTCTGGAATCAACATGCTCAACTGGCCCGACGAAACACAG 111252  
Qy 592 agcgtcaagttgaagatcgcgtctcctgtggtcgtgaaggtgtcttcgtcggaagacgag 651  
Db 111251 AGCTGCAAGTTGAAGATCGGCTCTCTGGGCTGTAAGTCTGCTGCGGAGAACGGCAGC 111192  
Qy 652 gcgagaggaagtcctcttgacccaacgaagacactgtgttaagtcacccgaggtggaatcgtg 711  
Db 111191 GCGAGAGAGAGTCCCTTGACACAGACGACACTGCTTCAGTCACCGGAGTGGAAATGCTG 111132  
Qy 712 gactcgcgagcccaacttctcgaatcgaactactaaggtacatagatgatacactctgag 771  
Db 111131 GACTCGGAGACCCCACTTGTGACAGACGACTACACGCTAATGAGTACACTCTGACG 111072  
Qy 772 gctcagcggtctctcattgatacagcgctcatcacaacccggtctcgtatgctc 831  
Db 111071 GCTCAGCGGCTCTCTCATGTATACGCGCTCATCTACACCCGCGCTCTGATCTGTC 111012  
Qy 832 atcctgacctctcaagctctctgctgctcccccacatggcgcggaagatcatgctc 891  
Db 111011 ATCTGCGCCTCTCAGCTTCTGTGCTGCTGCTCCACATGGGGCGGAGAAATCATGATC 110952  
Qy 892 aacggtctcatalcatcgtgtgactgcgcgtctcctcaatgtacttcgcacagctctcga 951  
Db 110951 AACGCGCTGTCTCATCATCATGTGATGCGCGCTCTCTCATGTACTTCGCCAGCTCTCCCA 110892  
Qy 952 gtgctgtcacaacatctcactgtg----- 978  
Db 110891 GTGCTGTCAACAATTAATCTCACTTGTGTGAGTGTCTCCCATTTTCATCACTTC 110832  
Qy 979 ----- 978  
Db 110831 AGTTCGAAAACTAATGATGTAATGATACCTTAACATGCGACGAATTAATCCGAATTAACGA 110772  
Qy 979 ----- gtaatcttcaagcaacagcctgtctatct 1010  
Db 110771 CCACTCACCGGCGATCTCTTAATTTGCAATATCTTCAAGCACGACCGCTGCTGTATCT 110712  
Qy 1011 gaagcttcacacatcgtgtgaaggtctgaatctctgacttgacaggaagacgaagacag 1070  
Db 110711 GAGGCTTCCACCATGCTGAGGTTCTGATCTGTACTGTGACCAAGCAAGCAAGAG 110652  
Qy 1071 ggcgttcgcgagagcgtgtgaagaaagctgtcgaacggtgacgtgtgacgtgtcgtc 1130  
Db 110651 GCGCTGCGGAGCGCTGAGAAAGCTGTGACAGGGGACCTGGGCAACGTGGCTGTGCT 110592  
Qy 1131 ctggtgttcaagcaacactgtgaggtcgaagcggaagaaacaaagatgtgacagca 1190  
Db 110591 CTGCGTGTTCAGCACACCGCGGATGTCAGGCGGAGAACCCAAAGAGATGTGACGAGCA 110532  
Qy 1191 cccgtagaagagcgagacagagaggttcaatgcgtcgtggaatcaaacacacagaggt 1250  
Db 110531 CCGCTACGAGAGGCGGACGAGAGTCCAGTCCCTGTGGCATYCAACACACCGAGGT 110472  
Qy 1251 gccggtgcgaagcgaacacaggttcgactgtgcgtgtcgtgtcgtgcgcgcgt 1310  
Db 110471 GCGGGGCGCAAGGCGCAACACAGTTCGACTGGGCGCTCTGGCCACGCGCGTGACGCGCAT 110412  
Qy 1311 ttccttcttctcctcagcctgtgctctctcactatctgtgcaatcaagtgtctcgtgtgag 1370  
Db 110411 TTCTTCTGTTTCTTCAAGCTGTGGCTTCTCTCATCTGTGGCATCAAGGTGCTCCGCTGTGAGG 110352  
Qy 1371 atgctcgaagactcaagccacatcccaagcgaatgtgacactctgaactaatttgcatt 1430  
Db 110351 ATGCTCGAGACTCAAGGCGCACATCCCAAGCCAGTGGCCACTCTGAATCTGATTTGCAATTT 110292  
Qy 1431 ggcattcatgtatlaattgtgtgtgcaacttaattlaattgaatgagacctgta 1490



/gene="CG11911"  
/product="CT37076"  
/db\_xref="FLYBASE:FBan0011911"  
/db\_xref="FLYBASE:FBgn0031249"  
complement(<13646..>14580)  
/gene="CG11911"  
/map="21B7-21B8"  
/db\_xref="FLYBASE:FBan0011911"  
/db\_xref="FLYBASE:FBgn0031249"  
complement(13743..14576)  
/gene="CG11911"  
/note="CG11911 gene product"  
/codon\_start=1  
/db\_xref="FLYBASE:FBan0011911"  
/db\_xref="FLYBASE:FBgn0031249"  
/protein\_id="AAFS1508.1"  
/db\_xref="GI:7296216"  
/translation="MKLITVTLVLAVALAAGAKLSDKLALVPSFATGEVINGTEAE  
PHSAPYIVSLATNYLKHSHICGFLINKDVIPTAHCISEPVGMSIAGLHTRAVEVE  
LTQOROVDFRHKYTGKGVDPYDIALHNGVIFENWOPATLPSEVQHEEHL  
GKNGOPKSYTFSCAKTQYTTOTLNEPEKELPEAPAESITCSSIQOSKSN  
GDSGGLVPEFTNAPSRLIGIVSMGYIPCLANPSTYTVASNTIDWITINQSRYTL  
F"  
complement(join(49281..49459,49730..50200,50254..50554,  
50786..>51230))  
/gene="CG11907"  
/product="CT3136"  
/db\_xref="FLYBASE:FBan0011907"  
/db\_xref="FLYBASE:FBgn0031250"  
complement(<49281..>51230)  
/gene="CG11907"  
/map="21C1-21C1"  
/db\_xref="FLYBASE:FBan0011907"  
/db\_xref="FLYBASE:FBgn0031250"  
complement(join(49281..49459,49730..50200,50254..50554,  
50786..51049))  
/gene="CG11907"  
/note="CG11907 gene product"  
/codon\_start=1  
/db\_xref="FLYBASE:FBan0011907"  
/db\_xref="FLYBASE:FBgn0031250"  
/protein\_id="AAFS1506.1"  
/db\_xref="GI:7296214"  
/translation="MPGSTDRRLVDALEDDEDELITGEGSFVPAFRPRODHNDROCLM  
ERODDEVLYVSNPEGRSLFTYLYFYLLIGSTKTRPMPFPTAEKRYKTRKNSINNT  
DLEBELTPLKSFCDLALATISGTFLDLNATFGHNSRTMLSTLMLLLEFV  
TTGFEVINTDTEQOFLILIVLLNISATMSGALYGAFLPSEFTTAVVSGA  
LGGLTALAFILVLAFTDGTNTAFIFIVIGSVLLICIVCYVILAKPFRYLYEGG  
DKKVIKAVSHNNGSAGEGLEPILROVMSKTYLAHSIALTYTTLTSYPAVTL  
MSEYGHSVTAMVMFMALNSGYTNILLMAPRKVKQHEKELASSIMAAALSCGMV  
GSLSLIVFVOML"  
join(66682..69037,75067..75450,76257..76302,76864..77132,  
77259..77995)  
/gene="a1"  
/note="Nucleotide sequence of the Celera sequence differs  
from the published sequence for this transcript."  
/product="CT12963"  
/db\_xref="FLYBASE:FBan0003935"  
/db\_xref="FLYBASE:FBgn0000061"  
<66682..>77995  
/gene="a1"  
/note="CG3935"  
/map="21C1-21C2"

Query Match 67.8%; Score 1044.2; DB 3; Length 302473;  
Best Local Similarity 77.2%; Pred. No. 5.2e-171;  
Matches 1518; Conservative 0; Mismatches 13; Indels 436; Gaps 3;

Qy 7 caccgaggtacatccgaacaagcgcgctgaacaatgaagcagcctcccaagataag 66  
Db 235713 GACGACCGTACATCCGAACAAGGCGGTGAACAATGACGACACATCCCAAGATAAG 235772

Qy 67 gcaccagttccgctcgtgactcactactgctgcaaatgctaattggagtgctctct 126

|||||  
Db 235773 GCACGAGTTCCGGHCTCGAGACTGCACACTAGCTGCAAAATGCTPATGGGATGCTT 235832  
Qy 127 atgggagctgaactccgtgcacagggcgaact----- 156  
Db 235833 ATGGGAGCTGACTTCCTGCTGACGAGGCGCACGTGAAGTGTATACGGCTCCGAGTTGCATA 235892  
Qy 157 -----gcacacgacgagcccaagaacgcaa 182  
Db 235893 AGCGGCTATATAGTGTGATCTCGAATCCCTTCACAGCCACCGCGAGCCCAAGAACGCCAA 235952  
Qy 183 lgtcaagggcccttgatcgcctccacacgcccgtgttcaagactaagcagaatgtgca 242  
Db 235953 TGTCAGAGCCCTTGATGCTGCTTCACAGCGGCGCTGTTCAAGAACTACGACAGATGTGCA 236012  
Qy 243 gccgagtgctccaaggaacccccacgaacgctgtccctgtgaatgtgtgtacactataga 302  
Db 236013 GCCGGTGTTCACAGGAAGACCCCGACGACGTCCTGGAAATGGTGTGCACCTACATGAGA 236072  
Qy 303 catcagcaggttgaacgagcagaagctgaacacccactgctgctgaatctc----- 351  
Db 236073 CATCAGCAGTTGAACGAGCAAGCTGACCCACTGCTGCTGAATCTGTAATGCCAGT 236132  
Qy 352 ----- 351  
Db 236133 CCTCTCCTTCGTCCTCTCCTCTGTCGAGTAGCTGAGAAACGATATTTCCGAGTTC 236192  
Qy 352 ----- 351  
Db 236193 CCTCAAGTTCACAACTCTGATAGCAATTTGAAAGCGGTGTATCTGAACTGAACTGA 236252  
Qy 352 ----- 351  
Db 236253 GTGTTGTTTGTGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 236312  
Qy 352 ----- 351  
Db 236313 GTGCATTTATGATATATCTGAATATCCATATCCATACCAAGAAAGCAATGGTGTTCAG 236372  
Qy 352 cgatgaggaagaagcagcgtgtgtgcaacogtcaaatatgacaatacaagcagatc 411  
Db 236373 CGATGGAAGACGAGAGGACGCGTGTGCAACCGTCAATATGACAAATCATACGCGAGATTC 236432  
Qy 412 acttgaagtcgaaggaagctctggaaccccccaatcaacacttcaagggagcgaaggt 471  
Db 236433 ACTTGAAGTCCAGGAGAGTCTGAGACCCCAATCAACACTTAAAGCGGAGCAAGGT 236492  
Qy 472 ggcctgaagccgaacaccaggtgacccctcaagccaagatggccaactccggtgtgagct 531  
Db 236493 GGCCTGATGCCGGAACCCAGGTGACCTCAACCCAGATGGCCACTTCCGGTGTGATGCT 236552  
Qy 532 ccaagcgtgtacaacgctactgcaactcaacatcaatgctcaacttggcccaagcagaag 591  
Db 236553 CCAGCGGTGTACACGCGCTACTAGCGAATCAACATGCTCAACTGCCCCACGACGAACGAG 236612  
Qy 592 agctgcaagttgaagatgctcctcgtggcgctgaaggtcgtcctcgtcggagaaagcagc 651  
Db 236613 AGCTGCAAGTTGAAGATGCGCTCTGGGGCCCTGAAGGTCTGCTCCGGAAGAACGGCACG 236672  
Qy 652 gcgaagagagatcccttgacaacacgaacagcgtgttcaagtaacacggaatggaataatc 711  
Db 236673 GCGAAGAGAGAGTCCCTTGACACGACGACGACCTGTTCACTGACGCGGAATGCGGTG 236732  
Qy 712 gactcgaagccacttgtcaagtaagtaagtaagtaagtaagtaagtaagtaagtaag 771  
Db 236733 GACTCGAGAGCCCACTTGTGACGTACGACTACTAGGCTACATGAGATGACTCTGACG 236792  
Qy 772 gctcagcagcgtcctcctcatatgtaacagcgctcatcacaaccccggtctcgtatcgtc 831  
Db 236793 GCTCAGCGGCGCTCTCATGATGACGCGCGCATCTACACACCCGCTCTGCACTCTC 236852  
Qy 832 atcctcgcctctcagcctctcgtgctcctcccaatcggcgaggaagatcatgct 891  
|||||

Db 236853 ATCTTGCCCTCTCAGCTTCTGCTGCTCCACCATGGGCGGAGAAATCATGATC 236912  
Qy 892 aacggcctgctcatcatcgtgacgcgcctctcactgactcgcagctctcgca 951  
Db 236913 AAGGGCTGCTCATCATCGATGAGCGCGCTCTCTCATGACTTCGGCAGCTCTGCCA 236972  
Qy 952 gtcgtcgaacaatactcactctgtg----- 978  
Db 236973 GTGCTGTCCAAACAATACTCCACTGTGGGAGTGTCCCATTTGCATTCAGTTC 237032  
Qy 979 ----- 978  
Db 237033 AGTTTGAATAAGTACTGATGATCCCTTAACATCGCAATATTACCAACTTACCGA 237092  
Qy 979 -----gtaattctctacagcacagcctgtctatct 1010  
Db 237093 CCACCTACCGGGCATTCCTTAATTTGAGTAATCTTCTTACAGCACGCGCTGCTATCT 237152  
Qy 1011 gagcgtctccacacatcgtcgaagttctagtctctgactcgtgcaagcaagaag 1070  
Db 237153 GAGCGTCTCCACCATGCTCAGAGTTCAGTTCGTACTGCGCACAGGCAAGCAAGAG 237212  
Qy 1071 ggcgcctgcggagggcgctggaagcgtcgtcgcgcgcgccttgagcgcgtgctgct 1130  
Db 237213 GCGCTGCGCGAGGCGCTGAGAAAGCTGTGCAAGGCGACCTGGGCACTGCTGCT 237272  
Qy 1131 ctgcgtgtctcagcacactgctgagctgcagcgcggaagaacaaagataggagca 1190  
Db 237273 CTCGGGTTCAGCACCGCGCGAGTGCAGGGGAGAAACACAGAGATGAGCGAGCA 237332  
Qy 1191 cccgtacagagagcggaagcagaagatccagtcgctggtgcatcaacacacagagt 1250  
Db 237333 CCGGTACGAGAGGCGGACAGAGAGTCCAGTCCGCTGGCATCAACACACCGAGGT 237392  
Qy 1251 ggcgggagcgaagggcgaagctgcagctgagcgtcgtcgcgcgcgcgtggaagcgt 1310  
Db 237393 GCGGCGCGCAAGGCCACACAGTTCGACTGGGCGTCTG6CAGCGCGCTGGAGCGCAT 237452  
Qy 1311 ttccctcgtctcctcagcctgagcctcctcctcctcctcctcagtcagtcgtcgtg 1370  
Db 237453 TTCTCTGTTTCCCTTCAGCTGAGCTGCGCTTCCTCATTTCTGGCATCAGGTGCGTGGT 237512  
Qy 1371 atgctcgaagcctcaagggcacaatcccaagcagtgcgactctgaactatgtgca 1430  
Db 237513 ATCTCGAGACTCAAGGCCACATCCCAAGCAGTGCACCTCGAATGATTTTGGCATTT 237572  
Qy 1431 ggcatttcagatgaatgaatgctgctgagcaactataatgaatgaatgaatgaat 1490  
Db 237573 GCGATTTCAATGATTTATATGTTGTCGAGACTTATATTTATATGATGAGACTGCTGA 237632  
Qy 1491 tggataaagagacactcgcgaatgctcgtcttaaaaaa 1537  
Db 237633 TGGAAATTAAGACCTCTGCCGATGCTGCTTCAACTAGAGGAA 237679

RESULT 9  
LOCUS Dm159B9T 299 bp DNA linear STS 17-APR-1996  
DEFINITION D. melanogaster STS determined from European Mapping Project  
ACCESSION 231885  
VERSION 231885.1 GI:469891  
KEYWORDS STS.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 299)  
AUTHORS European Drosophila Mapping Consortium.  
TITLE Direct Submission  
JOURNAL Submitted (08-APR-1994) Michael Ashburner, Department of Genetics,  
Downing St., Cambridge CB2 3EH, England

REFERENCE 2 (bases 1 to 299)  
AUTHORS European Drosophila Mapping Consortium.  
TITLE Direct Submission  
JOURNAL Submitted (15-APR-1996) Michael Ashburner, Department of Genetics,  
Downing St., Cambridge CB2 3EH, England  
REMARK Updated comments  
COMMENT STS\_name = Dm159B9T  
clone\_name = 159B9  
STS\_from\_promoter = T7  
vector\_class = cosmid, Lorist 6  
origin\_of\_clone = Oregon-R  
in\_situ\_site\_primary = 21C  
STS\_dbSTS\_AC = 4426  
BLAST\_program = BLASTN  
database\_searched = EMBL  
database\_version = 45.0 and updates till date\_of\_search  
date\_of\_search = 08-01-1996  
BLAST\_program = BLASTX  
database\_searched = SWISSPROT  
database\_version = 32.0  
date\_of\_search = 15-12-1995.  
location/Qualifiers  
FEATURES  
source 1..299  
/organism="Drosophila melanogaster"  
/strain="Oregon-R"  
/db\_xref="taxon:7227"  
/clone="159B9"  
BASE COUNT 59 a 77 c 89 g 74 t  
ORIGIN

Query Match 9.4%; Score 144.6; DB 11; Length 299;  
Best Local Similarity 97.4%; Pred. No. 2.3e-15;  
Matches 147; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 caagaggtacatccgaagaagcgcgctggaacatgacgcagctcccaagataag 66  
Db 168 CACACCCGTACATCCGAAMAAGGCGCGCTGAACAATGACGACGACTCCCAAGATTAAG 109  
Qy 67 gcaacagttcgcgtctcgtgacgcgcactcgtcgaatgctgaatggaatgctctt 126  
Db 108 GCACACGATTTCCGCTCTGACGTGCCACTACTCTGCAATGCTAATGGAGTCTTCTT 49  
Qy 127 atggggctgactcgtcgtccagcgccactg 157  
Db 48 ATGGGCTGACTCCGTCGACGCGCAGCGCCACTG 18

RESULT 10  
LOCUS GGA250362 1678 bp mRNA linear VRT 26-OCT-1999  
DEFINITION Gallus gallus mRNA for nicotinic acetylcholine receptor, beta-2  
subunit (nchr2b gene).  
ACCESSION AJ250362  
VERSION AJ250362.1 GI:6136929  
KEYWORDS beta-2 subunit; nchr2b gene; nicotinic acetylcholine receptor.  
SOURCE chicken.  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 1678)  
AUTHORS Nef, P., Oneyser, C., Alliod, C., Couturier, S. and Ballivet, M.  
TITLE Genes expressed in the brain define three distinct neuronal  
nicotinic acetylcholine receptors  
JOURNAL EMBO J. 7 (3), 595-601 (1988)  
MEDLINE 88283624  
REFERENCE 2 (bases 1 to 1678)  
AUTHORS Cooper, E., Couturier, S. and Ballivet, M.  
TITLE Pentameric structure and subunit stoichiometry of a neuronal  
nicotinic acetylcholine receptor  
JOURNAL Nature 350 (6315), 235-238 (1991)  
MEDLINE 91172320

REFERENCE	3 (bases 1 to 1678)
AUTHORS	Ballivet M.
TITLE	Direct Submission
JOURNAL	Submitted (25-OCT-1999) Ballivet M., Biochemistry, University of Geneva, 30 quai Ernest Ansermet, 1211 Geneva 4, SWITZERLAND
FEATURES	Location/Qualifiers
SOURCE	1. 1678
	/organism="Gallus gallus"
	/variety="white Leghorn"
	/db_xref="taxon:9031"
	/tissue_type="brain"
	/dev_stage="embryonic"
	/country="Switzerland"
	72. 125
sig_peptide	/gene="nachrb2"
	72. 1547
CDS	/gene="nachrb2"
	/codon_start=1
	/evidence=experimental
	/product="nicotinic acetylcholine receptor, beta-2 subunit"
	/protein_id="CAB59627.1"
	/db_xref="GI:6136930"
	/translation="MALTRVLCILALRSLCTDTERLVEYLIDPRYKILTRPATN GSLVLYQLMVSIAQLISVHERQIMNTNWLIOEMEDYRLTKPPEFDMKKYRLPS KHLMLDPVLYINNADGTEVSEFVSNVISTDGSIFWLPATVYSACKTEVKEHPEDQ NCTMRFSWYVINDREIDLVLSKSEVASLDDPTPSGEMDVALPGRNRPDOSTYDIT YDYLIRKPLFTVYINLIPCLILTSIALIVFYLPDGEKMTLCISVLAITVFLILI SKVLPSPSLDPLVGLKIMFTWLVTFPSIVTSVCLVHNRSPTHMPWVRLFLR KLALLFMKQPOONCARQRLRORRORERAAATFLRAGARACACVANGAKAEEL NGLRERQSGPDPBPARGCGLSEAEVSRVARIADHMRSEDDQSVSEDMKYVAAYIDRL FLMIFVFCVFTGVGMFLPLFQNTATNSLLDGLGGTPISK"
gene	72. 1547
	/gene="nachrb2"
	126. 1544
mat_peptide	/gene="nachrb2"
	/product="beta-2 subunit, nicotinic acetylcholine receptor"
BASE COUNT	320 a 591 c 480 g 287 t
ORIGIN	
Query Match	7.88; Score 120.4; DB 5; Length 1678;
Best Local Similarity	49.98; Pred. No. 3e-11;
Matches 394; Conservative	0; Mismatches 381; Indels 15; Gaps 3;
OY	297 catagacatgcagagatggaacgcgcaagctctgacacccacacgcgcgtgcgaatctccgat 356
Db	251 CATGACGCTGCACGACGCGGAGCGAGATCATACCAACCACTGTGGCTGACCCAGGAATG 310
Y	357 gagaagcagagagagcgctgtgtgcaacgcgtcaacaatgacacacatgacacacgcagatcaatt 416
Db	311 GGAAGACTATGCGCTCACCTGCGGAGACGAGACTTGACAACTGAAGAAGGTCGGCT 370
OY	417 gaagtccaagcgaggtctcggaccccccaaatcaactcttcaacggcggaaggtgacct 476
Db	371 GCCCTCCAAAGCACATCTGGCTGCTGATGTGTACTTACAAACAACGCGGAGGATGTA 430
OY	477 gatgcgcg-----aaaccacggtgacacctagccacgacgacgacgacgttgatgac 530
Db	431 CGAGGCTCTCTTCTTACTTCACAGCCGCTCATCTCTTACACGCGACACATCTTGTGGTGGC 490
OY	531 tccagcgctgtacacgacctactctgcgaactcaacatgctcgaactgcccccaagca 590
Db	491 CCCCAGCATACAAAGAGCGCGTGCAGAGATGAGAGTGAAGCATTTCCCTTCGACAAACA 550
OY	591 gaagctgaagtgtgaagatcgagctccctgggggcctgaaggtcgctctcggcggaagcgac 650
Db	551 GAATGCACCATGAAAGTTCCCTTCTGAGACTTACAGACGCGACCGAGATCGACTGTGCT 610
OY	651 ggcgagagagagagcccttgacacgagacgacctgttcagtcacacggaggtggaaatgct 710
Db	611 GAAGAGCGAGGTGGCCAGCCCTGGATGACTTACGCGCCACGCGGACAGTGGACATCTGGC 670

QY	711	ggaacgcgagagccacattgttcagccagagactactagcgctacacatgaggtacact-----	765
Db	671	GCTGCCGGGACGGGCCAACGAGAACCCGGACACTCCACACTAGCGAGACATCCTACGA	730
QY	766	-ctgacggcctcagcgcgctccctccatgtaacagcgctcatctacacacccgcctctg	824
Db	721	CTTCATCATCCGGCGCAAGCCGCTCTTCAGACACATCAACCTCATCACCCTGCATCCT	790
QY	825	catgtcatcctgtgcctctcagccttctgtctgctcctccacatgagcgcgagagaat	884
Db	791	CATCACCTCCCTGGGCAATCTCTCTTACTACCTCCCTCCGAC---TCGGAGAGAAAT	847
QY	885	catgtacagagcgccgcgcacatcatctgatatgcgcgccttccatcagtaactgtccagct	944
Db	848	GACGCTTGCAATCTCCGCTCTCTCTGCTGCCCTCACCTCTTCTCGCTCATCTCCAAAT	907
QY	945	ctcgacgaatgtctgtccacaatactaccactctgtgttaactctctacagaccagcctgt	1004
Db	908	CGTGCCGGCCACCTCGCTGAGAGTCCCGCTGCTGGGCAAGTACATCATGTTCCACATGT	967
QY	1005	gtaatcgcgcgtctccacatcgtctgaggttctagtctgtacctgtgcgcagagcaaga	1064
Db	968	GCTGTGTGACTTCTCTCATCTGCTCACCAGCGGTCTGCGCTGCTCAACGTGCACACCGCTCGCC	1027
QY	1065	caagagcgcg 1074	
Db	1028	CACCACTCAC 1037	
RESULT	11		
LOCUS	GG4250360	1608 bp	linear VRT 26-OCT-1999
DEFINITION	GG4250360	Gallus gallus mRNA for nicotinic acetylcholine receptor, alpha-2	
ACCESSION	AJ250360		
VERSION	AJ250360.1	GI:6136925	
KEYWORDS	alpha-2 subunit; nachra2 gene; nicotinic acetylcholine receptor.		
SOURCE	chicken.		
ORGANISM	Gallus gallus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.		
AUTHORS	1 (bases 1 to 1608)		
TITLE	Net.P., Oneysen, C., Alliod, C., Couturier, S. and Ballivet, M.		
JOURNAL	Genes expressed in the brain define three distinct neuronal		
MEDLINE	nicotinic acetylcholine receptors		
REFERENCE	EMBO J. 7 (3), 595-601 (1988)		
AUTHORS	2 (bases 1 to 1608)		
TITLE	Ballivet, M.		
JOURNAL	Direct Submission		
FEATURES	Submitted (25-OCT-1999) Ballivet M., Biochemistry, University of		
source	Geneva, 30 qual Ernest Ansermet, 1211 Geneva 4, SWITZERLAND		
	location/Qualifiers		
	1..1608		
	/organism="Gallus gallus"		
	/variety="white leghorn"		
	/db_xref="taxon:9031"		
	/tissue_type="brain"		
	/dev_stage="neonate"		
	/country="Switzerland"		
	7..75		
	/gene="nachra2"		
	7..1593		
	/gene="nachra2"		
	/codon_start=1		
	/evidence=experimental		
	/product="nicotinic acetylcholine receptor, alpha-2		
	subunit"		
	/protein_id="CAB59625.1"		
	/db_xref="GI:6136926"		
	/translation="MGWPCRSIIPLDWVCFTVLQATREOKOPHGAEDRLFKHLFTG		
	sig_peptide		
	CDS		





LOCUS	AR071400	2068 bp	DNA	linear	PAT 18-FEB-2000
DEFINITION	Sequence 1 from patent US 5910582.				
ACCESSION	AR071400				
VERSION	AR071400.1	GI:7222288			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 2068)				
TITLE	Elljott,K.J., Ellis,S.B. and Harpold,M.M.				
JOURNAL	Human neuronal nicotinic acetylcholine receptor compositions and				
FEATURES	methods employing same				
source	Patent: US 5910582 A 1 08-JUN-1999;				
	Location/Qualifiers				
	1..2068				
	/organism="unknown"				
BASE COUNT	410 a	634 c	582 g	440 t	2 others
ORIGIN					

Query Match	7.1%;	Score 109.4;	DB 6;	Length 2068;
Best Local Similarity	49.9%;	Pred. No. 2.3e-09;		
Matches 399; Conservative	0;	Mismatches 373;	Indels 27;	Gaps 4;

OY	297	catagaactcgagaggttgaagcgaaactgcgaaccccaactgcgttcgaactcctcgat	356
Db	450	CATGATGTGGATGTGAGAGAACCAATATGATGCCAACAGCTGTGGCTTAAACAGAGATG	509
OY	357	gagagacgagagagcgcgltgtgccaacgylacacatalatgacaacalcacgcagatcaatl	416
Db	510	GAGCAGCATACAAACCTGAGCCCTGGAAACCCCGCTGATTTTGGCAATCACATCATCTAGAGGT	569
OY	417	gaagtcacagcgaggtctggaaccccccaatacactctttcaac---ggcgagaaagctg	473
Db	570	CCCTTCTGAGATGATCTGGATCCCGACATTTTCTCTACAAACAANNTTGGGAGTTTGC	629
OY	474	cctgtagcggaaccccaagctgaacctgaagccaagatgycacatlcgcgttgatgcttc	533
Db	630	AGTGACCCCAATGACCAAGGCCCACTCTTCCACGGGCACTGTGCACTGGGTGCCCCC	689
OY	534	agccgtgacacagcgctactgccaactaaactagtcactgagccccacgaagaag	593
Db	690	GGCCATTACAAAGAGACTCTGTGAGATGAGCTGACCTTTCTCCCTTCGACACAGAGAA	749
OY	594	ctycaagttgaagatcgggtcctctggygcctgaagylcgylcctgcggaagcagcgagc	653
Db	750	CTGCAAGATGAAAGTTGGCTCTGTGACTTATGACAAGGCCAAGATGCG---ACCTGAGCA	806
OY	654	gagagagaagttccctctggaacccaagagacctggtttaagtlacacggaggtggaattg	713
Db	807	GATGGAGCAGATGTGGAGACTGGAAGAGCTACTGGAGAGAGGGCGAGTGGGCGATGTGAA	866
OY	714	ctcggagagcccaacttgttaagtcaggaactacgaagctacatacgag-----	759
Db	867	TGCCACGGGACCTTACACAGCAGCAAGAATGACAGCTGCGGCCGAGATCTACCCGAGGT	926
OY	760	----tacaactgacggtcagggcgctcctctcaatgtaacagcgccgtcatctacaac	815
Db	927	CACCTACGCGCTTGCATACCGGGCTCGCGCTCTTTCACACATCAACACTCATATCC	986
OY	816	cgagtcctgcatcgtcatcctgagccctcgaaccttcgagcttcgagctccccaatgagcg	875
Db	987	CTGCTGCTCATCTCTCTGCTACTGTGTGTCTTTCATACCTGCCCCCTCCGAC---TGGCG	1043
OY	876	cgagaagatcatgatacaacgycgtgcatcaatcgttgatcgcgccgtcctcattact	935
Db	1044	CGAGAAGATCAAGCTGTGCAATTCGGTGTGCTGTCACATCACCGTCTTCTCTGCTCAT	1103
OY	936	cgccagagctcctgcgcaggtgctgccaacaatactcaactgtggtgtaacttcttaagca	995
Db	1104	CACGTAGATCATCCCGTCCACCTCTGCTGGTCAATCCCGCTCATTCGGCGAGTACCTGCTGT	1163
OY	996	caagctgtctgatactgagcggtccacaacatcgtaggttcaggttcagttcttacctggaac	1055

Db 1154 CACCATGATCTTCTGTACCCCTGTCCATTCGATCACCCGCTTCTGTGCTCAATGTGGACCA 1223

Oy 1056 aggcagcacaaagagcgcg 1074

Db 1224 CCGCTCCCCCGACGCCAC 1242

RESULT 15

LOCUS	AR173184	2068 bp	DNA	linear	PAT 17-DEC-2001
DEFINITION	Sequence 1 from patent US 6303753.				
ACCESSION	AR173184				
VERSION	AR173184.1	GI:17912675			

ORGANISM	Unknown.
	Unclassified.
REFERENCE	1 (bases 1 to 2068)

JOURNAL	TITLE
US 6303753-A 16-OCT-2001;	Human neuronal nicotinic acetylcholine receptor compositions and methods employing same

source	1. .2068	/organism="unknown"	2 others
BASE COUNT	410 a	634 c	582 g
ORIGIN	440 t		

Query Match	7.1%	Score 109.4	DB 6	Length 2068
Best Local Similarity	49.98%	Pred. No. 2.3e-09		
Matches 399	Conservative	0	Mismatches 373	Indels 27
				Gaps 4

OY	297	catgacatcgacgaglttgaagcagcgcaagcttgacaccaccactgctgctgtaatctccgatg	356
Db	450	CATGATGTGGATGATGAAGAACCAAAATGATGACACCAACGTCCTGGCTTAAACAGAGTG	509
OY	357	gagagcagagagcgctgtggtgcacccgttaacatatgacaacatacgcagatcaactt	416
Db	510	GAGGACTACAAATCGGCGTGTGMAACCCCGTGAATTTTGGCAACATACATCTCTAGGGT	569
OY	417	gaagttcaagcgaggtctggaaccccccaatcacatcttcaac--ggcagcgaagtggtg	473
Db	570	CCCTTGTAGATGATGTGATGCCCGCATTTGTTCTTCAACAAANNNGGGAGATTGGC	629
OY	474	ccctgatgcggaacccaagttgaacctgaagcagatggcgaacttcggttgatgtcccc	533
Db	630	AGTGAACCCATGTGACCAAGGCCACCTCTTCTCCAGGGGACTGTGCATGTGGTCCCCC	689
OY	534	agccgtgtacacgcgctaccctgcgaactcaacatgctcaactggtccccaagcaagcagag	593
Db	650	GCCATCTACAGAGCTCTCTGACATCGACGTCACTTCTTCCCTTGTGACCAACAGAA	749
OY	594	ctgcaagttgaagtcgcgtctctctggtggtcctgaagtcgtctctgcccgaagcagcagc	653
Db	750	CTCGAAGATGAAGTTGGTGTCTCTGAGACTTATGACAAAGGCCAAGATG---ACCTGAGACA	806
OY	654	gagagagagatgcctttgacacagcagcagctggttcaagtaaccgagatggtgaatctgtga	713
Db	807	GATGGAGAGACTGTGGACCTGGAAGGACTACTGTGGAGAGCGGGAATGGGCCATGTCAA	866
OY	714	ctgcgagcccgaacttgtaagtaagacactacagctacatgagag-----	759
Db	867	TGCCACGGGCACTTACACACAGCAAGATGTGCACTCTGCGCGAGATGTACCCCGAGCT	926
OY	760	----taaccttgaacgctgaagcgcgtctctcatagtatacagcgcggtatctatacaccc	815
Db	927	CACCTAGCGCTTGTCATTCGGGGGCGTCCGCGCTTCTTACACATCAACCTCATCATGCC	986
OY	816	cgcgtctcgaatcgtaactcctggtcccttaagcctcttcgtgctgctccccaatggtgcg	875
Db	987	CTGCCTGTCTTCTTCGCTCCATCTGAGTGTGCTTCTTCACTGCCCCCTCCGAC---TGGG	1043



```

QY 876 cgagaagatcatgaaccaaggcctgctcatcatcgtgatcgccgcgcttcctcatgtaact 935
    |||||
Db 1044 CGAGAGATGACGCTGTGTCATTTCGTCCTGTCACTACCGCTTCCTGCTGCTCAT 1103
    |||||
QY 936 cgccagctcctcgcaagtgctgccaacaatactccacttggttaatctcttaagcac 995
    |||||
Db 1104 CACTGAGATCATCCCGTCCACCTGCTGTCATCCCGCTCATCGCGAGTACCTGCTGT 1163
    |||||
QY 996 cagcctgctgatatcgagcgcttcacacatcgtcgaggttctagtctgtaacctggccac 1055
    |||||
Db 1164 CACCATGATCTTCGTCAACCTGTCATGTCATCACCGTTCGTGCTCAATGTGACCA 1223
    |||||
QY 1056 aggcaagcacacaaggcgc 1074
    |||||
Db 1224 CCGCTCCCGCCAGCACCCAC 1242
    |||||

```

Search completed: July 18, 2002, 08:54:44  
Job time: 14268 sec





DR WPI: 2001-376906/40.  
DR P-PsDB: AAB86218.  
XX  
XX New nucleic acid encoding Drosophila acetylcholine receptor subunit,  
PT useful in screening for plant protection agents and pharmaceuticals -  
PT  
XX  
XX  
PS  
XX Claim 1; Page 10-13; 18pp; German.  
XX  
CC This invention describes a novel nucleic acid (I) which encodes a  
CC Drosophila melanogaster acetylcholine receptor beta subunit, Db3 which  
CC has insecticidal activity. (I), related host cells, polypeptides, ACR,  
CC antibodies, transgenic invertebrates and regulatory regions, are used to  
CC identify agents that modulate ACR activity, potentially useful as  
CC plant-protection agents (e.g. insecticides) or pharmaceuticals for human  
CC or veterinary medicine. This sequence encodes the Drosophila melanogaster  
CC acetylcholine receptor beta subunit described in the method of the  
CC invention.  
XX  
SO Sequence 1539 BP; 352 A; 476 C; 411 G; 300 T; 0 other;

Query Match	100.0%	Score 1539;	DB 22;	Length 1539;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1539; Conservative	0;	Mismatches	0;	Gaps 0

QY	1	atcgcgacgaggtgacatccgaatacaagagcgcgcttgacacatatgacgacgactcccaag	60
Db	1	atctcgacagaggtacataccgaatacaagagcgcgcttgacacatatgacgacgactcccaag	60
QY	61	ataaagacacacggtcttcggtctctgtagacacacatactctctcaaatgataatgtggatg	120
Db	61	ataaagacacacggtcttcggtctctgtagacacacatactctctcaaatgataatgtggatg	120
QY	121	ctcttatggggtgactacgtctgcgaagcgcgacactgccaacgcyggaaccccaagaagcgc	180
Db	121	ctctcttatggggtgactacgtctgcgaagcgcgacactgccaacgcyggaaccccaagaagcgc	180
QY	181	aatgtcaagggcccttgatctgccttccacgacgagcgctctgttcaacgaatacagacagcgatgtg	240
Db	181	aatgtcaagggcccttgatctgccttccacgacgagcgctctgttcaacgaatacagacagcgatgtg	240
QY	241	cagccgggtgttccaaagaaacccccacgaacgtgtcccttggaataatgtgtgtcaactacata	300
Db	241	cagccgggtgttccaaagaaacccccacgaacgtgtcccttggaataatgtgtgtcaactacata	300
QY	301	gacatcgacagagttgaacgcaagctgcacacacacactgtgtctgtgaattccgaatggaga	360
Db	301	gacatcgacagagttgaacgcaagctgcacacacacactgtgtctgtgaattccgaatggaga	360
QY	361	gagcgagagcggtgtgtgcaaacggtcacataatgacaacatcagcagatcatcttgaag	420
Db	361	gagcgagagcggtgtgtgcaaacggtcacataatgacaacatcagcagatcatcttgaag	420
QY	421	tccagcgaggttgtgaaacccccccaatatcaacacttcaataagcgagacgaatgtgacctgag	480
Db	421	tccagcgaggttgtgaaacccccccaatatcaacacttcaataagcgagacgaatgtgacctgag	480
QY	481	gccgaagacccaggtgacccctcagccacagaatggccacttccggttggatgctccagcgctg	540
Db	481	gccgaagacccaggtgacccctcagccacagaatggccacttccggttggatgctccagcgctg	540
QY	541	tacaaagccttaatggaataatcaaatgcttcaactgtgccccacagacagaagatctgcaag	600
Db	541	tacaaagccttaatggaataatcaaatgcttcaactgtgccccacagacagaagatctgcaag	600
QY	601	ttagaagatcggtctccgtgggcttgaaagtctgtcccttcggaagaaacgacagcgagaaagga	660
Db	601	ttagaagatcggtctccgtgggcttgaaagtctgtcccttcggaagaaacgacagcgagaaagga	660
QY	661	gagtcacctgacacagacgacctgtgttcaagtcacggaatgtggaaatcgttgacatcgga	720
Db	661	gagtcacctgacacagacgacctgtgttcaagtcacggaatgtggaaatcgttgacatcgga	720

QY	721	gccacatttgcagtcagtcagcctactaacgcttaatgagtaaacctctgaacgcttaacg	780
Db	721	gccacatttgcagtcagtcagcctactaacgcttaatgagtaaacctctgaacgcttaacg	780
QY	781	cgctccctccagtgacacagcgccatcctaaacacccgcgtctctgcatactgaatccctggcc	840
Db	781	cgctccctccagtgacacagcgccatcctaaacacccgcgtctctgcatactgaatccctggcc	840
QY	841	ctctcagcctctgctgctcctccccaatgagtcgagtcgagaaagatcatgatcaacgacctg	900
Db	841	ctctcagcctctgctgctcctccccaatgagtcgagtcgagaaagatcatgatcaacgacctg	900
QY	901	ctcatcatcgtgatcgcgcgcctctccccaatgactctgcaccagctccctgcagtgctctgc	960
Db	901	ctcatcatcgtgatcgcgcgcctctccccaatgactctgcaccagctccctgcagtgctctgc	960
QY	961	aacataatccactctgctggaatctcttaagacacagcctgcttatctagagcgtctcc	1020
Db	961	aacataatccactctgctggaatctcttaagacacagcctgcttatctagagcgtctcc	1020
QY	1021	acacatcgtcagagtgctctagctctctgtaacctgccccagacacaaagagagcgcctgcg	1080
Db	1021	acacatcgtcagagtgctctagctctctgtaacctgccccagacacaaagagagcgcctgcg	1080
QY	1081	gagcgctcgtagaaagactgctgcagcagccactgagtcgagtcgctgctcgtcgtctgc	1140
Db	1081	gagcgctcgtagaaagactgctgcagcagccactgagtcgagtcgctgctcgtcgtctgc	1140
QY	1141	agcacccactgctgcagtcgagc	1200
Db	1141	agcacccactgctgcagtcgagc	1200
QY	1201	gagcgctcgtagaaagactgctgcagcagccactgagtcgagtcgctgctcgtcgtctgc	1260
Db	1201	gagcgctcgtagaaagactgctgcagcagccactgagtcgagtcgctgctcgtcgtctgc	1260
QY	1261	aagcgcaaacagcttcgagctgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg	1320
Db	1261	aagcgcaaacagcttcgagctgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg	1320
QY	1321	tccttcacacgcgtgcctctccatctctgagccatagtgctccgttgtagagatgctcagaga	1380
Db	1321	tccttcacacgcgtgcctctccatctctgagccatagtgctccgttgtagagatgctcagaga	1380
QY	1381	ctcaagcgcaacatcccaagcagctgcagctgcagctgcagctgcagctgcagctgcagctgc	1440
Db	1381	ctcaagcgcaacatcccaagcagctgcagctgcagctgcagctgcagctgcagctgcagctgc	1440
QY	1441	gtatttaatgctgctgctgcagctataatatttaatgatgatgagacctgcgtatgagaaag	1500
Db	1441	gtatttaatgctgctgctgcagctataatatttaatgatgatgagacctgcgtatgagaaag	1500
QY	1501	gacctctgcgcgaatgctctgcttacaacaaaaaaaaaaaaa	1539
Db	1501	gacctctgcgcgaatgctctgcttacaacaaaaaaaaaaaaa	1539
RESULT 2			
ABL23317			
ID	ABL23317	standard; DNA; 1362 BP.	
XX	AC	ABL23317;	
XX	DT	26-MAR-2002 (first entry)	
DE	Drosophila melanogaster genomic polynucleotide SEQ ID NO 21424.		
KN	Drosophila: developmental biology; cell signalling; insecticide;		
KW	pharmaceutical; gene; ds.		
XX	Drosophila melanogaster.		
OS			
XX	WO200171042-A2.		

XX 27-SEP-2001.  
 PD  
 XX  
 PF 23-MAR-2001; 2001MO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 XX 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX MPI; 2001-656860/75.  
 DR  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Claim 1; SEQ ID NO 21424; 21bp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB101840-AB16175), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins  
 CC (AB07737-AB072072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SO Sequence 1362 BP; 294 A; 444 C; 374 G; 250 T; 0 other;

Query Match 87.8%; Score 1350.8; DB 23; Length 1362;  
 Best Local Similarity 99.5%; Pred. No. 7; 7e-306;  
 Matches 1355; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7 caagaggtatcatccgaacaaagcgcgctgaacaatgacagactcccaagataag 66  
 Db 1 caagacgtatcatccgaacaaagcgcgctgaacaatgacagactcccaagataag 60  
 QY 67 gcacaaattccggtctctgactgacactactgctgaatgataagggatctctc 126  
 Db 61 gcacaaattccggtctctgactgacactactgctgaatgataagggatctctc 120  
 QY 127 atggggtgacttcgtgacagggcgacactgacacgcgcgaccccaagaaacgcaatgctc 186  
 Db 121 atggggtgacttcgtgacagggcgacactgacacgcgcgaccccaagaaacgcaatgctc 180  
 QY 187 aaggccttgatagctccccaacgcgcgctgttcaagactacgaacagcgatgtgacacg 246  
 Db 181 aaggccttgatagctccccaacgcgcgctgttcaagactacgaacagcgatgtgacacg 240  
 QY 247 gtgttcaaggaacccccaagaaagtgtccctgtaaatgtgtgacactacataagacac 306  
 Db 241 gtgttcaaggaacccccaagaaagtgtccctgtaaatgtgtgacactacataagacac 300  
 QY 307 gacgaagtgaacggcaagctgacacacacactgtgctgtaattctcgaatgagaagacgag 366  
 Db 301 gacgaagtgaacggcaagctgacacacacactgtgctgtaattctcgaatgagaagacgag 360  
 QY 367 gaggcgctgtggaacacgctacaataatgacacacacagacagatcttgaattcagc 426  
 Db 361 gaggcgctgtggaacacgctacaataatgacacacacagacagatcttgaattcagc 420  
 QY 427 gaggctggagacccccaatcacacttcaacggcgacgaagtgccgtgataggccgaa 486  
 Db 421 gaggctggagacccccaatcacacttcaacggcgacgaagtgccgtgataggccgaa 480  
 QY 487 acccaagtgacacctcaacgaatgacacttcggtgatagtcctcgaagccgtgtacag 546  
 Db 481 acccaagtgacacctcaacgaatgacacttcggtgatagtcctcgaagccgtgtacag 540

Db 481 acccaagtgacacctcaacgaatgacacttcggtgatagtcctcgaagccgtgtacag 540  
 QY 547 gactactggaactcaacatgctaaactggtcccaagacaaagacgaatcgaattgaag 606  
 Db 541 gactactggaactcaacatgctaaactggtcccaagacaaagacgaatcgaattgaag 600  
 QY 607 atcggtcctggggcgctgaaagtgctctgcgcggaagaaacgacgacgaagagatcc 666  
 Db 601 atcggtcctggggcgctgaaagtgctctgcgcggaagaaacgacgacgaagagatcc 660  
 QY 667 ctgacacgaacgaactgtgttcagtcacgcgagtggaatactgtgactcgcgagccac 726  
 Db 661 ctgacacgaacgaactgtgttcagtcacgcgagtggaatactgtgactcgcgagccac 720  
 QY 727 ttgttcagtaaggaactactacgttaactgagtaactcttgaaaggtcagcggtcc 786  
 Db 721 ttgttcagtaaggaactactacgttaactgagtaactcttgaaaggtcagcggtcc 780  
 QY 787 tcaatgacacgcgcgtcatctacacacccgcgtctgcatcgtcatcctggccctca 846  
 Db 781 tcaatgacacgcgcgtcatctacacacccgcgtctgcatcgtcatcctggccctca 840  
 QY 847 gctctctgctgctccccaatgagcgcgagagaatcatgatacaagcgctgtcatc 906  
 Db 841 gctctctgctgctccccaatgagcgcgagagaatcatgatacaagcgctgtcatc 900  
 QY 907 atgtgtatcgccgctctctcaatgatacttgcagactctcgtcagatgctgtccaacat 966  
 Db 901 atgtgtatcgccgctctctcaatgatacttgcagactctcgtcagatgctgtccaacat 960  
 QY 967 actcaacttgtgtaactctctacagacacagcctgtgtatctgaaagctctcaacatc 1026  
 Db 961 actcaacttgtgtaactctctacagacacagcctgtgtatctgaaagctctcaacatc 1020  
 QY 1027 gtcgaggttctagtctctgacactggtgacacgaagacacaaagagcgctgcggaagcg 1086  
 Db 1021 gtcgaggttctagtctctgacactggtgacacgaagacacaaagagcgctgcggaagcg 1080  
 QY 1087 ctgagaaagctgtcgtgacacgagccctgtgacagtggtgtgtctgtgtgttcacac 1146  
 Db 1081 ctgagaaagctgtcgtgacacgagccctgtgacagtggtgtgtgtgttcacac 1140  
 QY 1147 actggtgagtcgacgagcggaagacacaaagagaatgacagacacccgttcgaagagcg 1206  
 Db 1141 actggtgagtcgacgagcggaagacacaaagagaatgacagacacccgttcgaagagcg 1200  
 QY 1207 gacgaagcagagatccagctcggtgacataacacacacgaagtgtccggcgccaagggc 1266  
 Db 1201 gacgaagcagagatccagctcggtgacataacacacacgaagtgtccggcgccaagggc 1260  
 QY 1267 aaccagttgactggggcgctgtggtgacacgcggtgacacgcatcttcttcgttcttc 1326  
 Db 1261 aaccagttgactggggcgctgtggtgacacgcggtgacacgcatcttcttcgttcttc 1320  
 QY 1327 agcctggtccttcctcatctcgtgacatcaggtgtccgtgtag 1368  
 Db 1321 agcctggtccttcctcatctcgtgacatcaggtgtccgtgtag 1362

RESULT 3  
 ABL23316/c  
 ID ABL23316 standard; DNA: 3798 BP.  
 XX  
 AC ABL23316;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 21421.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 XX pharmaceutical; gene; ds.  
 OS Drosophila melanogaster.

XX WO200171042-A2.  
 XX 27-SEP-2001.  
 XX 23-MAR-2001; 2001WO-US09231.  
 XX 23-MAR-2000; 2000US-191637P.  
 XX 11-JUL-2000; 2000US-0614150.  
 XX (PEKE ) PE CORP NY.  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 XX WPI; 2001-656860/75.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 XX Claim 1; SEQ ID NO 21421; 21pp + Sequence Listing; English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB163051), expressed DNA  
 CC sequences (AB161840-AB16175) and the encoded proteins  
 CC (AB57737-AB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 CC Sequence 3798 BP; 869 A; 1013 C; 1048 G; 868 T; 0 other;

Query Match 67.8%; Score 1044.2; DB 23; Length 3798;  
 Best Local Similarity 77.2%; Pred. No. 4.3e-234;  
 Matches 1518; Conservative 0; Mismatches 13; Indels 436; Gaps 3;

QY 7 caagaggttacatccgaacaaagcgcgctgaacaatgaagagcactcccaagataag 66  
 DB 2798 GACGACCGGATCCGCAAAAGGCGCTGCAATGACGACGACTCCCAAGATAAAG 2739  
 QY 67 gacacagttccggtccgctgactgactgactgactgactgactgactgactgact 126  
 DB 2738 GACACGATTCGCGCTCTGACTGCGACTGCGACTGCGACTGCGACTGCGACTGCG 2679  
 QY 127 atgggggtgactccgctgcagcgcgact----- 156  
 DB 2678 ATGGGGGTGACTCCGTCGACGCGCGCGCTGAACTGTTTACGGCTCCGAGTTGCATA 2619  
 QY 157 -----gcaacgcyggagcccaagaagcgcaa 182  
 DB 2618 ACCCGGTGAATGATGATCGAATCCCTCCAGCCGCGGAGCCCAAGAACCCCAA 2559  
 QY 183 tgtcaagggcctgactgactgactgactgactgactgactgactgactgactgactgact 242  
 DB 2558 TGTCAAGGCCCTGATCGCTCCACGCGCGCTGTTCAAGAACTACGAACGCAATGTGCA 2499  
 QY 243 ggcggtgttccaagaagccccaagaagtgctccctggaatgglygttactactataga 302  
 DB 2498 GCGGGTGTTCGAAGAACCCCAAGAACTGTCCTGGAATGGTGTGTCACCTACATAGA 2439  
 QY 303 celtgacgaggttgaagcggaagctgaaccaacacacacacacacacacacacacacac 351  
 DB 2438 CATTCGACGAGTTGAAGGCAAGCTGACACCACTGCTGCTGAATCTGTAAGTCCAGT 2379  
 QY 352 ----- 351  
 DB 2378 CCTCTCCTTCGCTCCTCTCTGCTCGAGTAGCTGAGAAACCAAGTAATTTCCGAGTTC 2319

QY 352 ----- 351  
 DB 2318 CCTCAAGTTCAACACTCTGATTAAGCAATTTGGAAAGCGTGTATCTGAAGTAAGTAA 2259  
 QY 352 ----- 351  
 DB 2258 GTGTGTGTTTGT 2199  
 QY 352 ----- 351  
 DB 2198 GTGCATTTTGAATATCTGAATATCCATATCACTAACCAAGCAATCGTGTGTCTAG 2139  
 QY 352 cgaatgagagagagagagcgctgtgtgcaacgctcaaatatgacacacacacacacacac 411  
 DB 2138 CGATGAGAGAGAGAGAGCGCGGTGTGCAACCGTCAATATGACAACTACACGAGATC 2079  
 QY 412 actttgaatcagaaggtctgtgaccccccaatatacactcttcaagcgagagaagt 471  
 DB 2078 ACTTTGAAGTCCAGCGAGGTCTGGACCCCAATATCACTCTTCAACGCGAGCAAGGT 2019  
 QY 472 ggcctgtatgccaagcccaagtgacctcaagcgaacgactccggtgtgtgact 531  
 DB 2018 GGCCTGATGCGCGAAACCCAGGTGACCTTCAGCCAGATGGCCACTTCGGTGTGCTCT 1959  
 QY 532 ccagccgtgtacaagcgctactgactgactgactgactgactgactgactgactgactgact 591  
 DB 1958 CCAGCCGTGTACAGCGGCTACTGCGAATCAACATGCTCAACTGCGGCCACGACAAGCAG 1899  
 QY 592 agctgaagtgaagatgagctgctccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 651  
 DB 1898 AGCTGAAATTTGAAGTCCGCTCTGGGCTGTGAAGGTGCTCTGCGGAGAAACGCGCAG 1839  
 QY 652 gcaagagagagatcccttaccacagacgacccgtgtgtgtgtgtgtgtgtgtgtgtgtgt 711  
 DB 1838 GCGAGAGAGAGATCCCTTACCAACGACGACCTGTTCATCACTACCGAGTGGAAATCGTG 1779  
 QY 712 gactcgagagccactgtgtcagtaagaactactgactgactgactgactgactgactgact 771  
 DB 1778 GACTCGAGGCCCACTTTGTGATGATGAGACTACTGACTGACTGACTGACTGACTGACTG 1719  
 QY 772 gctcagcggtgctctccatgtacaaagcgctcatataacacccggtgtgtgtgtgtgtgt 831  
 DB 1718 GCTCAACGGCGCTCTCCATGTAACAGCGCGTCTATCTAACAACCCGCTCTGACATCGTC 1659  
 QY 832 atccgtgaccttcaagccttccgt 891  
 DB 1658 ATCTGTGCGCTCTGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1599  
 QY 892 aacggtgtctatcatcgtgtatgcgcgcttccatcgttactcgtccagctcctgtgca 951  
 DB 1598 AAGGCGCTCTCATCATCGTATCGCGCGCTCTCATGTATGATGCGCCAGCTCTGCGCA 1539  
 QY 952 gtgctgtccaaatactcactgtgt----- 978  
 DB 1538 GTGCTGTCCAAATAATCTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1479  
 QY 978 ----- 978  
 DB 1478 AGTTCTGAANAATAAGTGTAGTACCTTAACCTTAACCTTAACCTTAACCTTAACCTTA 1419  
 QY 979 -----gtaactcttaagaacagcgtgtgtatct 1010  
 DB 1418 CCACCTACCGGCGCATTTCTATTAATTGACATTAATCTCTACACACACAGCTCTGTATCT 1359  
 QY 1011 gagggtgtccacatctgtgtgaggttctgtatctgttaccctggcacaagaagacaagaag 1070  
 DB 1358 GAGCGTCTCCACATCTGT 1299  
 QY 1071 ggcgctgcgagagcgctgagaagctgtgtcagcgagcctgggagagctgtgtgtgtgtgtgt 1130  
 DB 1298 GCGCGTCCGAGGCGCTGATGAAGAGCTGTGCAAGGACCTGTGGAGCTGTGTGTGTGTGT 1239  
 QY 1131 ctggtgttcaagcaaccatgctgagctgagcgagagagagcaaccagaagatgagagagca 1190

RESULT 4  
 ABL05370/c  
 ID ABL05370 standard; cDNA; 7041 BP.  
 XX  
 AC ABL05370;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 10592.  
 XX  
 KW Drosophila: developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO2001/1042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 XX  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI: 2001-656860/75.  
 DR P-PSDB; ABB61267.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Claim 1; SEQ ID NO 10592; 21np + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention  
 CC is useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins  
 CC (AB57137-ABR72072).

Query Match	35.7%;	Score 549.4;	DB 23;	Length 7041;
Best Local Similarity	97.2%;	Pred. No. 1.9e-118;		
Matches 559;	Conservative 0;	Mismatches 16;	Indels 0;	Gaps 0;

RESULT	5
ABL23312/c	
ID	ABL23312 standard; DNA; 3764 BP.
XX	
AC	ABL23312;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster genomic polynucleotide seq ID NO 21409.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide
KW	pharmaceutical; gene; ds.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US09231.
XX	

PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 PA (PEKE ) PE CORP NY.  
 XX  
 XX  
 PI Venter JC, Adams M, Li PMD, Myers EM;  
 XX  
 XX WPI; 2001-656860/75.  
 DR  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 XX Claim 1: SEQ ID NO 21409; 21bp + Sequence Listing; English.  
 PS  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins  
 CC (AB101840-AB116175).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pot\_sequences.  
 CC  
 XX  
 XX Sequence 3764 BP; 943 A; 972 C; 960 G; 889 T; 0 other;

Query Match 17.2%; Score 265.4; DB 23; Length 3764;  
 Best Local Similarity 82.2%; Pred. No. 3.3e-52;  
 Matches 346; Conservative 0; Mismatches 11; Indels 64; Gaps 1;

QY 7 caagagggatcacccgaagagggcgctgacgaatgacgagcagctcccaagataag 66  
 DB 485 CACACACCGTACATCCGAACAAAGCGCGCTGAACAATGACGAGACTCCCAAGATAAAG 426  
 QY 67 gacacagttccgctcgtcgtgactgcactactgtgtaataatgagtgatctctt 126  
 DB 425 GCACGAGTTCCGGTCTCTGAGTGCACATCTCTGCAATATGTAATGGAGATCTTCTT 366  
 QY 127 atggagcgtgactcgtgcgaagcgccact----- 156  
 DB 365 ATGGGCTGACTTCGCTGCAGCGGCACACGTAACTTTACGGCTCCGAGTTGCATA 306  
 QY 157 -----ggcaccggagaccccaagaagccaa 182  
 DB 305 AGCCGCTGATAGTGTGATTCGAATCCCTCCAGCCGCGGACCCCAAGAACCCCAA 246  
 QY 183 tctcaagcccttgatcgctcccaagcgcgctgttcaacgaactacgaacgagtgtga 242  
 DB 245 TGTCAAGGGCGCTGATCGCTCCAGCGCGGCTTCACGAATATAGACACGATGTGCA 186  
 QY 243 gccggtgttccaaagaaccccaagacgtgtccctggaatgtgtgtcaactacataga 302  
 DB 185 GCGGCTGTCCAAAGGAACCCCAAGACGTCCTGGAATGTGTGTACACTACATAGTA 126  
 QY 303 catcgaagcgttgaacgaagcgaaccccaactgtgtgtcgttaactcgaatggagaga 362  
 DB 125 CATCGAGAGTTGAACGGCAGCTGACCACTGCTGTGCTGAATCTGTAAGTCCAGT 66  
 QY 363 c 363  
 DB 65 C 65

RESULT 6  
 AAV12199  
 ID AAV12199 standard; cDNA; 2277 BP.  
 XX  
 AC AAV12199;  
 XX

DT 14-MAY-1998 (first entry)  
 XX  
 XX Human neuronal nicotinic acetylcholine receptor alpha-2 subunit cDNA.  
 DE  
 XX  
 XX Human: neuronal nicotinic acetylcholine receptor: alpha-2 subunit;  
 KW Brain tissue; screening; NACHR; antibody; ds.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 FH Key location/Qualifiers  
 FT CDS 166..1755  
 FT /tag= a  
 FT /product= "neuronal nicotinic acetylcholine receptor  
 FT alpha-2 subunit"  
 XX  
 XX  
 PN M09420617-A2.  
 XX  
 PD 15-SEP-1994.  
 XX  
 PF 08-MAR-1994; 94WO-US02447.  
 PF  
 PF 08-MAR-1993; 93US-0028031.  
 PR  
 XX (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.  
 PA (SIBI-) SIBIA NEUROSCIENCES INC.  
 XX  
 PI Elliott KJ, Ellis SB, Harpold MM;  
 DR WPI; 1994-303024/37.  
 DR P-PSDB; AAW44155.  
 XX  
 XX Human neuronal nicotinic acetylcholine receptor subunits and DNA -  
 PT also transformed cells useful for screening cpds. which modulate  
 PT activity of the receptor  
 XX  
 XX Disclosure; Page 67-68; 99pp; English.  
 PS  
 XX  
 CC The present sequence encodes a human neuronal nicotinic acetylcholine  
 CC receptor (NACHR) subunit. The cells expressing the alpha and/or beta  
 CC NACHR subunits may be used in a method of screening compounds to  
 CC identify any which modulate the activity of human neuronal NACHR.  
 CC Subunit specific antibodies may be used to monitor the distribution  
 CC and expression density of various subunits in normal vs diseased brain  
 CC tissues. Testing of single receptor subunits or specific receptor  
 CC subunit combinations with a variety of potential agonists or antagonists  
 CC provides information with respect to the function and activity of the  
 CC individual subunits and should lead to the identification and design of  
 CC compounds that are capable of very specific interaction with one or  
 CC more receptor subtypes. The resulting drugs should exhibit fewer  
 CC unwanted side effects than drugs identified e.g. screening with cells  
 CC that express a variety of subtypes.  
 CC  
 XX  
 XX Sequence 2277 BP; 451 A; 695 C; 643 G; 486 T; 2 other;

Query Match 6.9%; Score 106.8; DB 15; Length 2277;  
 Best Local Similarity 49.9%; Pred. No. 3.1e-15;  
 Matches 400; Conservative 0; Mismatches 372; Indels 30; Gaps 4;

QY 297 catagacatcgacagatgaacggaagctgaccaccactgtgtgaatcctcgatg 356  
 DB 450 catcgtgtgtgatgagaagaacaaatgtagaccacacgctgtgctaacaagagtg 509  
 QY 357 gagagacgagggagcggtgtgtgcaacgcgcacaaatagtacacatracgacatcatt 416  
 DB 510 gagcactacaacatcgctgtgaacccgcgtgatttggcaacatcactctcctcagggt 569  
 QY 417 gaatcgaagcgtgtgacccccaatcacaactctcacaagcgacgaag----- 470  
 DB 570 cccctctgagatgatctgtatcccgacacttctctcacaacatgacgagatggaggt 629  
 QY 471 tggcctgatgcgcgaacccaggtgaccctcagcgaatggtgacctcgttggatgcc 530  
 II III I III III III I I III I I III III



```

Db 630 tgcagtgacccacatgacaaagccacacctctctccacagggcactgtgcctgtgccc 689
Oy 531 tccagcgtgtacacagcgctactgtgaactcaactgtcactgtgccccacgaagaac 590
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 650 cccggccatctacaagaagctctctgcagacgcagctcaactctctcccttcgacagca 749
Oy 591 gagctgcaagtgtgaacatcgctcctctgggacctgagaggtctctctcgcggaagaagc 650
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 750 gaactgcagaatgaaattgtgtctctctgtgactatgacaagccaatgcg---acttga 806
Oy 651 ggcgagagagagatgcctcttgacacgacgacactgtgtcaatcacgagtggaatcgt 710
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 807 gcagatgtgagcagactgtggaacctgaagactactgtgagagcgcgagtgggccatcgt 866
Oy 711 ggactcgcgagccactctgtcagtcagactactagcgctacatgagag----- 759
Db 867 caatgtccacgggcaccttcaacaagaagaagacgactgtgcgcgagatctaccgca 926
Oy 760 -----facactctgaagcgtcagcgcgctccctccatctgtacacgcgctcatcagac 812
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 927 cgtcacctacagccttcgcacatccgcgctgcgcgctctctctcacaacatcactatcat 986
Oy 813 acccggtctctgcacatcgtcactcctcagcctctcagcctctgtgcctcccaatgag 872
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 987 cccctgcctgcacatctctcctcactcgtgtgtgtctctctcactgcgcctccgcac---tg 1043
Oy 873 cggcgagaagaatcatgatcaacgagcctgtcatctcgtatgtatgcgcgctctctcatgta 932
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1044 cggcgagaagaatcagcgtgtgatttgcgtgtgtgtcgtcgtacacacgctctctcgtcgt 1103
Oy 933 ctgcgccagactcctgcacagctgtgtccacaatatcactctgtgtgaatctctacag 992
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1104 catcacatgatacatccctccacactcgtgtcatccgcgtcactcgcgagatctcgt 1163
Oy 993 caccagcctgtctatctgtgagcgtctccacaatcgtcgaaggttactgttctgtacctg 1052
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1164 gtccacatgatactctgcacacctgtgtcatcgtcatcaccgctctcgtcgaatgtgca 1223
Oy 1053 cacaagcagaagcacaagagcgcg 1074
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1224 ccaacgctcccccagacccac 1245

```

```

RESULT 7
AAQ90387
ID AAQ90387 standard; cDNA: 2277 BP.
Xx
Xx AAQ90387;
Xx AC
Xx CT
Xx 30-NOV-1995 (first entry)
Xx
Xx Alpha 2 subunit of a human neuronal nicotinic acetylcholine receptor.
Xx DE
Xx Human nMACHr; neuronal nicotinic acetylcholine receptor;
Xx KW
Xx neurotransmitter; ss.
Xx
Xx Homo sapiens.
Xx OS
Xx
Xx Key Location/Qualifiers
Xx FH 166..1755
Xx FT /tag= a
Xx
Xx MO9513299-A.
Xx
Xx PD 18-MAY-1995.
Xx
Xx PF 08-NOV-1994; 94MO-US12859.
Xx
Xx PR 08-NOV-1993; 93US-0149503.
Xx
Xx (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.
Xx
Xx Elliott KJ, Ellis SB, Harpold MM;
Xx

```

```

Xx WPI: 1995-194036/25.
DR P-PSDB: AAR73966.
Xx
Xx New human neuronal nicotinic acetylcholine receptor alpha2 subunit DNA
Xx PT used to develop prods. for detection, diagnosis and therapy and for
Xx PT modulating activity
Xx
Xx Claim 2; Page 43-46; 54pp; English.
Xx
Xx DNA encoding the human nMACHr alpha2 subunit was isolated from a
Xx human thalamus tissue cDNA library using corresp. rat cDNA. The
Xx insert of one clone obtd. was ligated with the insert of another
Xx clone to generate a full-length alpha 2 subunit cDNA. The DNA can be
Xx used to identify function nMACHrs. Cells contg. the DNA can be used
Xx for screening to identify cpds. which modulate the activity of human
Xx nMACHrs. The human nMACHr alpha 2 subunit can be used to product
Xx antibodies which can be used in immunohistochemistry, diagnosis and
Xx therapy. The nucleic acids can be used for analysing disease states
Xx and creating animal models.
Xx
Xx Sequence 2277 BP; 451 A; 696 C; 643 G; 487 T; 0 other:
Xx
Xx
Xx Query Match 6.9%; Score 106.8; DB 16; Length 2277;
Xx Best Local Similarity 49.9%; Pred. No. 3.1e-15;
Xx Matches 400; Conservative 0; Mismatches 372; Indels 30; Gaps 4;
Xx
Xx 297 catagacatcgacgagttgtgaacggaagctgacacccactgtgtcgtgaatctcgaatg 356
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 450 catcagatgtgataggaagaacaaatgatgaccccaacgtctgtcaaacagagatg 509
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 357 gagaagcagagagcgcggtgtgcaacgctcaacaatgatcaacatcagcagatcattt 416
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 510 gagcgaccacaacacgtgcgtgtgaacccgcgtgatttggcaacatcactctcaaggt 569
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 417 gaagtcacgagaggtctgtgacccccaatcactcttcaacgycgacgaag----- 470
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 570 cccctctgagatgatctgtgatcccgacatgttctctacaacaatgacagtggagatt 629
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 471 tggcctgttgccgaacaccaggtgacacctcaagccaagatggccactcgggtgtgagc 530
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 630 tgcagtgacccacatgacaaagccacacctctctccacagggcaactgtgcctgtgccc 689
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 531 tccagcgtgtatacagcgctactgtgaaactcaactgtcactatctgtccccacgaagaac 590
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 650 cccggccatctacaagaagctctctgcagacgcagctcaactctctcccttcgacagca 749
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 591 gagctgcaagtgtgaagatcggtctcctgggacctggaaggtgtctgcgcggaagcagac 650
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 750 gaactgcagaatgaaattgtgtctctgtgactatgacaagccaatgcg---acttga 806
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 651 ggcgagagagagatgcctcttgacacgacgacactgtgtcaatccagctgtggaatcgt 710
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 807 gcagatgtgagcagactgtggaacctgaagactactgtgagagcgcgagtgggccatcgt 866
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 711 ggactcgcgagccactctgtcagtcagactactagcgctacatgagag----- 759
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 867 caatgtccacgggcaccttcaacaagaagaagacgactgtgcgcgagatctaccgca 926
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 760 -----facactctgaagcgtcagcgcgctccctccatctgtacacgcgctcatcagac 812
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 927 cgtcacctacagccttcgcacatccgcgctgcgcgctctctcacaacatcactatcat 986
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 813 acccggtctctgcacatcgtcactcctcagcctctcagcctctgtgcctccccaatgag 872
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 987 cccctgcctgcacatctcctcgtcactcgtgtgtcttcttaacctgcctccgcac---tg 1043
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 873 cggcgagaagaatcatgatcaacgagcctgtcatctatcgtgtatgcgcgctctctcatgta 932
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 1044 cggcgagaagaatcagcgtgtgatttgcgtgtgtgtcgtcgtacacacgctctcgtcgtcgt 1103
Xx ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Xx 933 ctgcgccagctcctgcagatgctgtccaacatctcacttgtgtgaatctctacag 992

```

```

Db 1104 caccacgagatcatcccgccacacctgctgcatccgctcatcgcgagtaacctgct 1163
OY 993 caccagcctctgtatctgagcgctccaccatcgctgaggtctagtctgacctgac 1052
Db 1164 gtccacatgatctctgtcacacctgctccatcgctcatcaccgctctgtcacaatgca 1223
OY 1053 cacaagcagaacacagaagcgac 1074
Db 1224 ccacgcctcccccaagcaccac 1245

RESULT 8
AAT48235
ID AAT48235 standard; DNA: 2664 BP.
AC AAT48235;
XX
XX 09-APR-1997 (first entry)
DT
XX
XX Neuronal nicotinic acetylcholine receptor alpha-2 subunit DNA.
XX
XX Neuronal nicotinic acetylcholine receptor; nAChR; neurotransmitter;
XX ligand-gated receptor; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH 555..2141
XX FT CDS /*tag= a
XX
XX WO9641876-A1.
XX
XX 27-DEC-1996.
XX
XX 07-JUN-1996; 96MO-US09775;
XX
XX 07-JUN-1995; 95OS-0484722.
XX
XX (SIBI-) SIBIA NEUROSCIENCES INC.
XX
XX PI Elliott KJ, Harpold MM;
XX
XX WPI; 1997-065463/06.
XX
XX DR P-PSDB; AAM09021.
XX
XX Nucleic acids encoding nicotinic acetylcholine receptor sub-units -
XX used in screening to determine the effect of drugs on the receptor
XX
XX Disclosure; Page 49-52; 108bp; English.
XX
XX A DNA sequence (AAT48235) codes for the alpha-2 subunit (AAM09021) of
XX the human neuronal nicotinic acetylcholine receptor (nAChR). Host
XX cells, esp. mammalian cells or amphibian oocytes, carrying alpha-2
XX nucleic acids, opt. in combination with other alpha and/or beta
XX subunit nucleic acids (see also AAT48232-34, AAT48236-41), express
XX recombinant nAChR subunits useful for identifying cpds. that
XX modulate the activity of human nAChRs.
XX
XX Sequence 2664 BP; 518 A; 815 C; 743 G; 588 T; 0 other;

Query Match 6.9%; Score 106.8; DB 18; Length 2664;
Best Local Similarity 49.9%; Pred. No. 3.2e-15;
Matches 400; Conservative 0; Mismatches 372; Indels 30; Gaps 4;

OY 297 catagacatcgacgagtgtaagcgcaagctgaccaccactgtgtgctgaatctcgatg 356
Db 839 catcgatgtgatgagagaacaaatgatgaccaccaagctgtgctaaacagagagt 898
OY 357 gagaagagagagcgcgctgtggaacacgtacacatgatgacacacgcagatcattt 416
Db 899 gagcgactacaactgcctggaaccccgatcttgacacacatcattctcagggt 958
```

```

OY 417 gaatcagcgaagtcctggaaccccccaatcaactcttcaacgcgagcaag----- 470
Db 959 ccctctcgagatgatctcggtctcccgacattgtctctcaaaaacagatggggagtt 1018
OY 471 tggcttgatgacggaacccaggtgaacctcagcacaagatgacacttcggtgatgac 530
Db 1019 tgcagtgaaccacatgaccaaagcccaactctcttccacagggacgtgtgacatgggtgc 1078
OY 531 tcaagcgtgttacacgagcctgacactggaactcaacatgctaaatggtcccaacagaaga 590
Db 1079 cccggccatctacaagaagctccctgcagacatgacgtcaactcttcccttcgacaga 1138
OY 591 gaactgcaagtgtgaagaatcgctctggggcctggaaggttcctctccggaagaagcgcac 650
Db 1139 gaactgcaagaatgaaagtgttgctctcctggaacttatagacaagccaagatcg---acctgga 1195
OY 651 ggcgagagagagatcccttgacaacgaagacactggttcaagtcacacggatggtggaatcgt 710
Db 1196 gcagatggaagcaagactgtgacactgaagactacttggagagcggcgtggtgacatcgt 1255
OY 711 ggactcgagcccaacttgtcagtcagactactacggtacatgag----- 759
Db 1256 caatgcacgggacacttaacacagaagaagtacagactgtctgcgcgaagatcaccacca 1315
OY 760 -----tacactctgacggtcagcggtcctctccatgtacacggtcatalctaacac 812
Db 1316 cgtcacactaagcctctgtaacccggtcgcgtctctcttcaacatacactcaatcat 1375
OY 813 acccggtctctgatactgatacctctgacctctcagacctctgtgctgctcccaatggg 872
Db 1376 cccctgcctgctatctccctgacctcactgtgctgttcttactctccctccac--tg 1432
OY 873 cggcgaagaagatgatcaacagcgctgcatcatcatgtagtcgcccctctctcaatga 932
Db 1433 cggcgaagaagatcatcgatcgatcttcggtgctgctgtaacacacgcttctctgctgt 1492
OY 933 ctgcgccagctctgccaagtgtgttccacaataactcaactgtgtgtaacttcttacaag 992
Db 1493 catcactgagatcatcccgccacacctgctgtaacccgctacatcgcgagtaacctgct 1552
OY 993 caccagcctctgtatctgagcgctccacacatcgctcagaggtctgattctgtaacctgac 1052
Db 1553 gtccacatgatctctgtccacctgtccatcagctacatcaccgtctctgtaaatgta 1612
OY 1053 cacaagcagaacacagaagcgac 1074
Db 1613 ccacgcctcccccaagcaccac 1634

RESULT 9
AAV12196
ID AAV12196 standard; cDNA: 2363 BP.
XX
XX AAV12196;
XX
XX 14-MAY-1998 (first entry)
DT
XX
XX Human neuronal nicotinic acetylcholine receptor alpha-4 subunit cDNA.
XX
XX Human; neuronal nicotinic acetylcholine receptor; alpha-4 subunit;
XX brain tissue; screening; NChR; antibody; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH 173..2056
XX FT CDS /*tag= a
XX FT /product= "neuronal nicotinic acetylcholine receptor
XX PN alpha-4 subunit"
XX
XX WO9420617-A2.
```



CC A DNA sequence (AA048237) codes for the alpha-4 subunit (AA090023) of  
 CC the human neuronal nicotinic acetylcholine receptor (nAChR). Host  
 CC cells, esp. mammalian cells or amphibian oocytes, carrying alpha-4  
 CC nucleic acids, opt. in combination with other alpha and/or beta  
 CC subunit nucleic acids (see also AA048232-36, AA048238-41), express  
 CC recombinant nAChR subunits useful for identifying cpds. that  
 CC modulate the activity of human nAChRs.

XX Sequence 3496 BP; 590 A; 1215 C; 1084 G; 606 T; 1 other;

#### Query Match

Best Local Similarity 6.8%; Score 104.6; DB 18; Length 3496;

Matches 420; Conservative 0; Mismatches 409; Indels 30; Gaps 4;

QY 297 catagacatcgacgaagttgaagcgaagctgacccacacgtgtgtgtatccgatg 356  
 DB 450 catgacgctggtatgataagaacacagatgatgacacgaacgtatgtgagcagagatg 509  
 QY 357 gaagagcagagagcgctgtgtgcaacgltcacataatgacaacatcagcagatcatt 416  
 DB 510 gcaagactacaagctgctgcgtggaacacagctgactatgagatgtacacccatccgat 569  
 QY 417 gaagtcacagcaggtcttgagacccccaatcacactcttaacagcgagaggtgctt 476  
 DB 570 cccctcagagctatctgtggtgagcagacatcgtctcttaacaacatgtgacggagact 629  
 QY 477 gatggtcgga-----accaggtgacccctcagcagcagatgacacttccggtggtgctc 530  
 DB 630 cgtggtac 689  
 QY 531 tccagcgtgtgatacagcgtctactgcaactcaacatgtctaacatgtgccccacagaaga 590  
 DB 690 cccgacatttaacagactctctcagcagcatcgtacacacacacacacacacacacacac 749  
 QY 591 gacttgaagttgaagatcgctctcgtggtggtcgtgaaggtctctcgcggaagaagcag 650  
 DB 750 gaactgcacacatgaatctggtctctgtgacactacagacaagcacaagatcgactgtgaa 809  
 QY 651 ggtgagagagagatcccttgcacacacacacacacacacacacacacacacacacacac 710  
 DB 810 catgacacag---ccgctgtgacacacacacacacacacacacacacacacacacacac 866  
 QY 711 ggaactgcgagacacacttgtcagtaagatactacacacacacacacacacacacacac 759  
 DB 867 ggtgacgctgtggtgac 926  
 QY 760 -----tacacttgaacggtcagcggtctctctcactacacacacacacacacacac 812  
 DB 927 catcactatgacttgcgtac 986  
 QY 813 acccggtctcgtacatgac 872  
 DB 987 cccctgctcgtacatctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1043  
 QY 873 cgtgaggaagatataatgaacgctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 932  
 DB 1044 tggcaggaagatataatgaacgctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1103  
 QY 933 ctgcgcacgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 992  
 DB 1104 catcagcagatatacctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1163  
 QY 993 caacagcgtcgtatctatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1052  
 DB 1164 gtctacacatgattcttgcacacacacacacacacacacacacacacacacacacac 1223  
 QY 1053 caacagcagacacagagcgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1112  
 DB 1224 cgt 1283  
 QY 1113 ggtcagcgtgtcgtcgtc 1131

DB 1284 cgtgcacgctgtcctc 1302

#### RESULT 11

ID AAS91552 standard; cDNA; 1809 BP.

XX AAS91552;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #27356.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR P-PSDB: ABG27365.

PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity

XX Claim 1; SEQ ID NO 27356; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (II) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1809 BP; 332 A; 665 C; 481 G; 331 T; 0 other;

#### Query Match

Best Local Similarity 6.8%; Score 104.4; DB 23; Length 1809;

Matches 411; Conservative 0; Mismatches 421; Indels 24; Gaps 3;

QY 297 catagacatcgacgaagttgaagcgaagctgacccacacgtgtgtgtatccgatg 356  
 DB 144 catgacgctggtatgataagaacacagatgatgacacgaacgtatgtgagcagagatg 203

Oy	357	gagagcagcagagagcgcgctgagcacaacggtacaaatgtgacaacatcacgcagatcattt	416
Db	204	gcaacgtattacaagctgcgtgcctgggagccacagctgactatgtgaatgtcaacctccatccgat	263
Oy	417	gaagttccagcgaaggttctgtgaccccccacaatcacactcttcaacccgagcgaagctgacct	476
Db	264	cccttcgcagatcatctatctgctggcgagcagatcatgtctctcttaacaagaatgctgaacggggatt	323
Oy	477	gatyggccgaaa-----accaggttgacccctcaagccacagatgtgccaattccggttgatgcc	530
Db	324	cgcgtgacacccacacccacgacaaagggcccaactgtgtccaaatgaacgggcgggtgtcagttgacctcc	383
Oy	511	tccagccggttgacacggcgctactatgtgaaactcaacatgctcaactgtgccccacagacaa	590
Db	384	cccgagccatttacaagagctccctgcagcagcatcgagcgtacacgtctctcttcccttcgagaccaga	443
Oy	591	gagctgcgaagtttgaagatcgcgtctctctggggc-----ctgaaggtcgtacct	635
Db	444	gaactgtcacacatgaagatctgcgtctctgcagcttaacagacgaagccaaagatcgacctgtgaa	503
Oy	636	gcccggagaaacggcagcggcgaagagagatgcctcttgacacacgaagacgtgttcagtcacc	695
Db	504	catgcacaacgcgcgttgagacacagcttggaactcttgaggagatggcgaagtgtgcatcgttga	563
Oy	696	ggagttgggaaatcgttgagctcgcgcgagcccaactctgtgcagtcgaagactactaagtgatcat	755
Db	564	tgcctgtgggcacactacacacacccagaaatgagatgtctgtgcgaagatctaacccggat	623
Oy	756	ggagttacactctgacggctctacgcgcgtctccatcattgatacggccgttcattacacacac	815
Db	624	cacctatgtccttgatcaccggcgtgcgcgtctcttcttacaacatcaactcatcatcc	683
Oy	816	cgcgtctctgatatgtatcctctgtgcctctcaagcctctctgtgcctccccaatgagcgg	875
Db	684	ctgcctgtcactatctcctgcctcacacgtgtcgtctcttcaactctgcctccga---gtgtgg	740
Oy	876	cgaagaatcatgtgtacaagcgctctgcattacatcgtgatacgcgcgccttccattactt	935
Db	741	cgaagaagatcacgctgtggaatctccgtgtcgtcgtcgtcgtccacgcgtcttccgtcgtcat	800
Oy	936	cgcaccagctccctgcagtgctgttccaaatactaccacttgtgtatccttctaagacac	995
Db	801	caccgagatcatcccgctcacaccctcaactgtgtacatcccatccatccgcgagttacctgtgt	860
Oy	996	cagcctcctgtatctatgtagcgtctccacaacatcgtcgaagttctagtctgtacctggcac	1055
Db	861	caccatgatctctgtcacccctgtgcatacgtgtacaaagcttctcgtctaaagcttgacaa	920
Oy	1056	agggcaacgacaaaggggcgtcgcgcggaagggcgtcgaagaagcgtgcgtcaacgggcacgtgg	1115
Db	921	ccgctgcgcacgcagcagcaacacccatgtcccaactcgtggtaacgaaggtcttccctgtgaatcgt	980
Oy	1116	caagtgtgctgtcgtc 1131	
Db	981	gccacgcctgtcctc 996	
RESULT 12			
AAT359528			
ID AAT359528 standard; DNA: 2082 BP.			
AAT359528:			
XX	AC		
XX	AT		
XX	05-OCF-1997	(first entry)	
XX	DE	Alpha4 subunit of mutant nACHR (TCC 248 TTC).	
XX	XX		
XX	nACHR; mutation: autosomal dominant nocturnal frontal lobe epilepsy;		
KW	ADNFLE; neuronal acetylcholine receptor, ss.		
XX	OS		
XX	Homo sapiens.		

FT	Key	Location/Qualifiers
FT	misc_RNA	902..1168
FT		/tag= a
FT		/note= "ion channel"
FT	mutation	1016..1018
FT		/tag= b
FT		/note= "TCC 248 TTC is the ADNFLE-
FT		inducing mutation (see AAU59527)."
XX		
XX	PN	AU9656247-A.
XX	PD	09-JAN-1997.
XX		
XX	PF	28-JUN-1996; 96AU-0056247.
XX	PR	28-JUN-1995; 95AU-0003840.
XX		
XX	PA	(UYBO-) UNIV BONN.
XX	PA	(UYME ) UNIV MELBOURNE.
XX	PA	(WOME-) WOMEN'S & CHILDREN'S HOSPITAL.
XX		
PI	Berkovic SF, Mulley JC, Phillips HA, Propping PJ,	
PI	Scheffer IE, Steinleinok, Sutherland GR, Wallace RH;	
XX		
DR	WP1: 1997-100506/10.	
DR	P-PSDB; AAM11825.	

PT CHRNA4 gene encodes neuronal nicotinic acetylcholine receptor  
PT fragment - used in diagnosis of autosomal dominant nocturnal frontal  
PT lobe epilepsy  
XX  
PS Claim 6, Fig 13; 20pp; English.

CC The DNA sequence of the normal nacrR is given in AAT59527,  
CC the DNA sequence of the ADNFLF-inducing mutant is given in AAT59528  
CC Mutations at codon 248, which have the effect of replacing  
CC serine by another amino acid (e.g., phenylalanine) in the sixth  
CC amino acid of the transmembrane domain 2 (M2) of the alpha4  
CC subunit of nacrR have been found to be associated with ADNFLF.  
CC The primers given in AAT59529 and AAT59530 were used in the  
CC amplification of part of exon 5.

Sequence 2082 BP; 367 A; 767 C; 592 G; 356 T; 0 other;

Query Match	6.7%;	Score 102.8;	DB 18;	length 2082;
Best Local Similarity	47.9%;	Pred. No. 2.6e-14;		
Matches 410; Conservative	0;	Mismatches 422;	Indels 24;	Gaps 3;

Oy	297	caagacatcgaagatttgaaacggcaagcttgaacaccactgcgtggttgatcttcgaatg	356
Db	397	cattgaacgtgtatagaagaaccagatgatatgaacacgaagctatggtgtgaagcagaagt	456
Oy	357	gagaagacgaagagcgcgtgtgtgcaaccgttcaacatgatgaacaatcaacgacatcatt	416
Db	457	gcaagcatatacaagcttgcgtctgtgaccacagctgatactatgaatgatgacccctccacat	516
Oy	417	gaatgccagcgaagttcttgagaccccccaatatcaactcttcaacgctgacgaagaatgtgct	476
Db	517	ccctctcgagctatctctgtggcgccgagatcgtctcttacaacaatgtctgaagggactt	576
Oy	477	gattgacgaa-----acccaagtgaacctcaacacagatgagccaattccggtgatgcc	530
Db	577	cgcgctaccaccaacttgacaaagagcccaactgttcttcaagacggtgctgtagatgatacc	636
Oy	531	tccagccgattgacacgacctctctcgaaactacaatcatgctcaactgcgcccacgaagaag	590
Db	637	cccgccattacaagagctctctcgacatctgcagatcgaagctcaactctctcccttcgacaga	696
Oy	591	gaagctgcaagttagaatcgctctctctgggc-----ctgaagtgcgtcct	655
Db	697	gaactgacccttgaaattcgctctctcttgactacgaagaagccaagatgcacctggtgaa	756

QY	636	gccgggaaagagcagcgcgaggaagaaagcccttcacacgcagccgctgtccagtcac	695
Db	757	catggacagcgcggtgagccagctgactcttcgtggagagatggcggtgtgtatcgtg	816
QY	696	gagatgggaaatctgtggaactcgctgcgagcccaattgtcagtcaggaactacagctacat	755
Db	817	tgcctgttggcaccttaaacacacaggaagatcgatgtgtgtgcgagatctacccggacat	876
QY	756	ggagtaacacttcagcgctcagcggtgcgcctcccaatgatacggcgctcatctaaacc	815
Db	877	cacctaagccttctcatccggcgccctgcgcctctctctacacatacaatcatatccc	936
QY	816	cgcgtccatcgtcatccatcccttgagccctcctaagcctctgtgctgcctcccaatgtggcg	875
Db	937	ctgcctgtcatctctctgcctccacacgtgtgtgtgtctctacatccctgcctccga---gtgtgg	993
QY	876	cgagaaagatgatcaaacgcgctgtgtacatctgtgatacggcgccctccatcatctact	935
Db	994	cggaaagatcatagcgtgtgtactctgtgtcgtcgtgtcgtccacgcgtctctccgctgtcat	1053
QY	936	cgcccaactctctgcagttgtgtgcacaaatatactcaacttgggtgaatcttctacagac	995
Db	1054	caccgagatcatccctgcctccacccctcaactgttcatccacatcatcgcgggtacgtgtgt	1113
QY	996	cagcctgtctatctgtagaggtctccacacatctgttgaggtgtcagttctgtatccctggcac	1055
Db	1114	caccatgatcttcgtacacccctgtgcacatgtcatcaaggtccttcgtgtccaagcttcacaa	1173
QY	1056	aggcaaaagcaaaagcgcccttcgcggaagcgctgtgaaagagctgtctgcacggtgcactgg	1115
Db	1174	cgcctgcgcacagcaagcaacacatgtccacactcgtgtgaacgaggtcttcctctgaaatcgt	1233
QY	1116	caagctgtgtctgtctc	1131
Db	1234	gccacgtctgtctctc	1249

RESULT	13
AAT59527	
ID	AAT59527 standard; DNA; 2082 BP.
XX	
AC	AAT59527;
XX	
DT	05-OCT-1997 (first entry)
XX	
DE	Alpha4 subunit of normal nAChR.
XX	
OS	nAChR; mutation; autosomal dominant nocturnal frontal lobe epilepsy. ADNFLE; neuronal acetylcholine receptor; ss.
XX	
MS	Homo sapiens.
XX	
FH	Key
FT	misc_RNA
FT	Location/Qualifiers
FT	902..1168
FT	/tag= a
FT	/note= "ion channel"
FT	1016..1018
FT	/tag= b
FT	/note= "TCC 248 TTC is the ADNFLE-
XX	inducing mutation (see AAT59528) "
PN	AU956247-A.
XX	
PD	09-JAN-1997.
XX	
PF	28-JUN-1996; 96AU-0056247.
XX	
PR	28-JUN-1996; 95AU-0003840.
XX	
PA	(UYBO-) UNIV BONN
PA	(UYME) UNIV MELBOURNE.
PA	(WOME-) WOMEN'S & CHILDREN'S HOSPITAL.
XX	

PI	Berkovits SF,	Mulley JC,	Phillips HA,	Propping PJ;
PI	Scheffer IE,	Steinleink,	Sutherland GR,	Wallace RH;
XX	WPI;	1997-100506/10.		
DR	P-PSDB;	AAM1824.		
XX				
XX	CHRNA4 gene encodes neuronal nicotinic acetyl:choline receptor			
PT	fragment - used in diagnosis of autosomal dominant nocturnal frontal			
PT	lobe epilepsy			
XX				
PS	Claim 5; Fig 13; 20pp; English.			
XX				
CC	The DNA sequence of the normal nAChR is given in AAT59527,			
CC	the DNA sequence of the ADNFLE-inducing mutant is given in AAT59528.			
CC	Mutations at codon 248, which have the effect of replacing			
CC	serine by another amino acid (e.g. phenylalanine) in the sixth			
CC	amino acid of the transmembrane domain 2 (M2) of the alpha4			
CC	subunit of nAChR have been found to be associated with ADNFLE.			
CC	The primers given in AAT59529 and AAT59530 were used in the			
CC	amplification of part of exon 5.			
XX				
XX				
S0	Sequence 2082 BP; 367 A; 768 C; 592 G; 355 T; 0 other;			

Query Match	6.6%;	Score 101.2;	DB 18;	Length 2082;
Best Local Similarity	47.8%;	Pred. No. 6.1e-14;		
Matches 409; Conservative	0;	Mismatches 423;	Indels 24;	Gaps 3;

QY	297	catgaagatcgaagatcttgaaacgcaagctctgaacacccactctgtgcgaatcccgatg	356
Db	397	cattgaaatgtgtagaagaagaacacgaatgtatgacacgaacgatactggtggaagcggatg	456
QY	357	gaagagacgaggaagcgcgtgtggtgcaaccgtcaacaataatgacaacatcaacgcatcatctt	416
Db	457	gcaacgaactaacaagctcgcgcggagaccagctgactctgagaaatgtacatccatccatccgat	516
QY	417	gaagttcagcgaggtctgtgaaccccccaaatgcacatcttcaacgagcgacgaagttgcgtc	476
Db	517	ccccctcgaagctatccctgtgcgcgcgcggaacatctgtccctctcaacaacatgtcgtgaagggatct	576
QY	477	gattgacgaa-----accacaggtgaacactcgaacgaatgagcgaacttcgcgtgtgatgccc	530
Db	577	cgcgtatcaaccacacttgacaagaagccacactgtttccaatgacggtgcgggtgtgaatgcatcc	636
QY	531	tccaagccggtgtaacgagcgtcactctgcgaactcaacatgtccaatctgcccacgaagaagca	590
Db	637	ccggccacatttaaaagatcccttgcacgaacatcgaaacttctcccttcacgaacacga	696
QY	591	gagctgcaagtgtgaagatcgcgtccctcgagggc-----ctgaagatcgtctc	635
Db	697	gaactgcgaaccatgaatcgcgtccctcgtgcacatcgaaagcgcaagatcgcacctgtgtaa	756
QY	636	gcccggaaacggaacgacgaggaagagagatcccttgacaacgaacgaactggtctcaagtcac	695
Db	757	catgcaacaacgcgcgttgacaacagcttggaactcttggaagagttgacgagttggtatcgttga	816
QY	696	ggaggtggaatatgttgagactcgcgcgaagcccaactttgttcaatgaagactacgtatcat	755
Db	817	tgcgcgtgggaacacttaacaacacgaagaagtacgagtgcgtgcgcgaagatctaacccgaatc	876
QY	756	ggagtaaacctctgaacggtctgaacgagcgtccctccatgtacaacgagcgtcatcttaacaac	815
Db	877	caactatgctctgtcatccgcgcgtcgcgtctcttcaacacatccatccatcatcc	936
QY	816	cgcgtctcgtatctgtcatcccttgcccttcacagcttcttgctcgtctcccaatgagggcg	875
Db	937	ctgcgtcgtcatctctcgtcctaacgcgtgtgttctcttaacctgcgcctccga---gttgagg	993
QY	876	cgagaagatcatatcaaacgcctctgtcaatcaatcgttgatcgcgcgccttccatctatctt	935
Db	994	cggagaagatcaacgctgtgatactccgcgtgtgctgtgcgcacaacgcttccctcgtcgtcat	1053
QY	936	cgcacagctctgcgcagtgctgttccaacaatactccacttgtgtgaatctctcaagcac	995







---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2002, 10:24:32 ; Search time 2887.37 Seconds  
(without alignments)  
11154.074 Million cell updates/sec

Title: US-09-732-680a-1

Perfect score: 1539

Sequence: 1 attcgacacgaggtacatc.....cttacaacaaaaaaaaa 1539

Scoring table:

OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

ord size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_bt: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vi: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htgo\_inv: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
--------	-------	-------	-------	--------	----	----	-------------

1	1539	100.0	1539	6	AX113259	Sequence
2	1421	79.3	1421	2	AY005148	Drosophila
3	1181	76.7	1334	3	DME318761	AC018761 Drosophila
4	628	40.8	85095	3	AC004573	AC004573 Drosophila
5	628	40.8	142257	2	AC017491	AC017491 Drosophila
6	628	40.8	183726	3	AC008002	AC008002 Drosophila
7	628	40.8	302473	3	AE003589	AE003589 Drosophila
8	507	32.9	1646	3	AY061622	AY061622 Drosophila
9	144	9.4	299	11	DM159B9T	231885 D. melanoga
10	10	1.6	74135	2	AC101533	AC101533 Mus muscu
11	11	1.6	77575	2	AC096600	AC096600 Rattus no
12	12	1.6	161361	2	AC090939	AC090939 Homo sapi
13	13	1.6	178622	2	AC026020	AC026020 Homo sapi
14	14	1.6	191494	2	AL591488	AL591488 Homo sapi
15	15	1.6	239135	2	AL591607	AL591607 Mus muscu
16	16	1.5	20149	2	HSL129H7B	Z54146 Human DNA s
17	17	1.5	62537	2	AC101100	AC101100 Mus muscu
18	18	1.5	110000	2	AL136360-2	Continuation (3 of
19	19	1.5	146407	2	AC004953	AC004953 Homo sapi
20	20	1.5	154998	2	AC109787	AC109787 Bos tauru
21	21	1.5	173512	2	AC068565	AC068565 Homo sapi
22	22	1.5	186026	2	AC099785	AC099785 Homo sapi
23	23	1.5	195773	2	AC073134	AC073134 Homo sapi
24	24	1.5	200622	2	AL158068	AL158068 Homo sapi
25	25	1.5	202375	2	AC109798	AC109798 Bos tauru
26	26	1.4	1135	8	AF134302	AF134302 Arabidops
27	27	1.4	1488	8	AY070424	AY070424 Arabidops
28	28	1.4	133260	2	AC024523	AC024523 Homo sapi
29	29	1.4	143769	2	AC004914	AC004914 Homo sapi
30	30	1.4	154673	2	AC079472	AC079472 Mus muscu
31	31	1.4	168463	2	AC021463	AC021463 Homo sapi
32	32	1.4	180844	10	AL606536	AL606536 Mouse DNA
33	33	1.4	187746	2	AC110003	AC110003 Homo sapi
34	34	1.4	189781	2	AC027506	AC027506 Homo sapi
35	35	1.4	217831	2	AC108162	AC108162 Homo sapi
36	36	1.4	230254	2	AC074163	AC074163 Mus muscu
37	37	1.4	239057	2	AP001578	AP001578 Homo sapi
38	38	1.4	254278	2	AC023836	AC023836 Mus muscu
39	39	1.4	293043	2	AC073707	AC073707 Mus muscu
40	40	1.4	581	9	AF029670	AF029670 Homo sapi
41	41	1.4	1632	9	AF267864	AF267864 Homo sapi
42	42	1.4	1731	10	AY012159	AY012159 Mus muscu
43	43	1.4	2030	3	AY051926	AY051926 Drosophila
44	44	1.4	2688	9	HSM800644	AL096749 Homo sapi
45	45	1.4	6656	6	AX281211	AX281211 Sequence
46	46	1.4	6656	6	AX356484	AX356484 Sequence
47	47	1.4	23777	2	AC106371	AC106371 Rattus no
48	48	1.4	41887	6	AX059526	AX059526 Sequence
49	49	1.4	57506	2	AC109246	AC109246 Mus muscu
50	50	1.4	68759	9	AL136111	AL136111 Human DNA
51	51	1.4	88848	8	T24H324	AF075598 Arabidops
52	52	1.4	90248	8	AC026480	AC026480 Arabidops
53	53	1.4	94536	9	AC010587	AC010587 Homo sapi
54	54	1.4	117621	9	AC003950	AC003950 Homo sapi
55	55	1.4	124456	2	AC017830	AC017830 Drosophila
56	56	1.4	136037	9	AC004104	AC004104 Homo sapi
57	57	1.4	136723	9	AC074000	AC074000 Homo sapi
58	58	1.4	139771	9	AC007425	AC007425 Homo sapi
59	59	1.4	141048	2	AC007644	AC007644 Homo sapi
60	60	1.4	143720	10	AC087332	AC087332 Mus muscu
61	61	1.4	148927	2	AC087127	AC087127 Mus muscu
62	62	1.4	152558	2	AC013487	AC013487 Homo sapi
63	63	1.4	154998	2	AC109787	AC109787 Bos tauru
64	64	1.4	158698	2	AC018625	AC018625 Homo sapi
65	65	1.4	161798	2	AC068062	AC068062 Homo sapi
66	66	1.4	164773	2	AC091543	AC091543 Felis cat
67	67	1.4	167912	2	AC007650	AC007650 Drosophila
68	68	1.4	171087	9	AC0073584	AC0073584 Homo sapi
69	69	1.4	173352	2	AC015543	AC015543 Homo sapi
70	70	1.4	173475	2	AC021864	AC021864 Homo sapi
71	71	1.4	175064	9	AC093848	AC093848 Homo sapi
72	72	1.4	175384	9	AL451103	AL451103 Human DNA
73	73	1.4	177298	2	AC068751	AC068751 Homo sapi

74	21	1.4	177673	2	AL365182	147	20	1.3	2596	5	DRE315610	AJ315610	Danio rer
C 75	21	1.4	178982	2	AC105656	148	20	1.3	2636	9	AK021422	AK021422	Homo sapi
76	21	1.4	179454	3	AC007646	149	20	1.3	2680	9	BC001336	BC001336	Homo sapi
C 77	21	1.4	188701	2	AC025242	150	20	1.3	3090	3	AY051596	AY051596	Drosophila
C 78	21	1.4	190664	2	AC090984	C 151	20	1.3	3535	8	YSPCTL	D55675	Yeast gene
C 79	21	1.4	190848	2	AC097310	C 152	20	1.3	4415	6	AX281491	AX281491	Sequence
C 80	21	1.4	198022	8	ATC8R1V1	C 153	20	1.3	4415	6	AX346808	AX346808	Sequence
C 81	21	1.4	203291	2	AC020998	C 154	20	1.3	4415	6	AX348855	AX348855	Sequence
C 82	21	1.4	208418	2	AE003701	C 155	20	1.3	4498	10	MMU11494	U11494	Mus muscu
C 83	21	1.4	209003	2	AC003724	C 156	20	1.3	4955	20	AB036063	AB036063	Homo sapi
C 84	21	1.4	209685	2	AC026231	C 157	20	1.3	5068	9	HSB801925	AL137265	Homo sapi
C 85	21	1.4	216877	2	AC105707	C 158	20	1.3	5236	6	AX345252	AX345252	Sequence
C 86	21	1.4	220491	2	AL445206	C 159	20	1.3	5340	6	AX345137	AX345137	Sequence
C 87	21	1.4	256417	2	AC079422	C 160	20	1.3	5817	9	AB028975	AB028975	Homo sapi
C 88	21	1.4	322436	2	AC074169	C 161	20	1.3	5825	6	AX323670	AX323670	Sequence
C 89	20	1.3	290	6	AX256044	C 162	20	1.3	5947	9	AF006513	AF006513	Homo sapi
C 90	20	1.3	290	6	AX198718	C 163	20	1.3	5964	6	AX344763	AX344763	Sequence
C 91	20	1.3	290	6	AX209250	C 164	20	1.3	6208	6	AX345096	AX345096	Sequence
C 92	20	1.3	290	6	AX209660	C 165	20	1.3	6352	6	AX345807	AX345807	Sequence
C 93	20	1.3	364	11	G49426	C 166	20	1.3	6361	6	AX251331	AX251331	Sequence
C 94	20	1.3	403	6	AX072066	C 167	20	1.3	6361	6	AX346736	AX346736	Sequence
C 95	20	1.3	420	6	IROEST042	C 168	20	1.3	6478	6	AX277961	AX277961	Sequence
C 96	20	1.3	502	6	AX255948	C 169	20	1.3	6478	6	AX323656	AX323656	Sequence
C 97	20	1.3	564	8	AF279895	C 170	20	1.3	6478	6	AX344793	AX344793	Sequence
C 98	20	1.3	590	8	SP1ESB1A	C 171	20	1.3	6608	6	AX346937	AX346937	Sequence
C 99	20	1.3	792	9	BC005905	C 172	20	1.3	6626	6	AX251568	AX251568	Sequence
C 100	20	1.3	797	8	AF244121	C 173	20	1.3	6716	6	AX346685	AX346685	Sequence
C 101	20	1.3	806	8	AF059294	C 174	20	1.3	6949	9	HSB297557	AJ297557	Homo sapi
C 102	20	1.3	892	8	AK000610	C 175	20	1.3	8129	6	AX251521	AX251521	Sequence
C 103	20	1.3	914	8	AY050945	C 176	20	1.3	8136	6	AX345457	AX345457	Sequence
C 104	20	1.3	917	10	BC009166	C 177	20	1.3	8136	6	AX348343	AX348343	Sequence
C 105	20	1.3	985	10	BC015940	C 178	20	1.3	8547	6	AX251816	AX251816	Sequence
C 106	20	1.3	1018	3	PFU41269	C 179	20	1.3	8547	6	AX344200	AX344200	Sequence
C 107	20	1.3	1040	10	BC006787	C 180	20	1.3	8547	6	AX348603	AX348603	Sequence
C 108	20	1.3	1138	8	BC000863	C 181	20	1.3	8801	6	AX277981	AX277981	Sequence
C 109	20	1.3	1194	9	HSU94317	C 182	20	1.3	8801	6	AX323672	AX323672	Sequence
C 110	20	1.3	1201	9	EC005398	C 183	20	1.3	8801	6	AX346643	AX346643	Sequence
C 111	20	1.3	1203	9	BC004240	C 184	20	1.3	9106	6	AX281327	AX281327	Sequence
C 112	20	1.3	1234	9	HSB322580	C 185	20	1.3	9838	6	AX281304	AX281304	Sequence
C 113	20	1.3	1241	10	AF146524	C 186	20	1.3	9838	6	AX345481	AX345481	Sequence
C 114	20	1.3	1242	10	MMU250491	C 187	20	1.3	10279	6	AX279995	AX279995	Sequence
C 115	20	1.3	1247	8	AY056230	C 188	20	1.3	10279	6	AX346492	AX346492	Sequence
C 116	20	1.3	1254	8	AY061131	C 189	20	1.3	10279	6	AX356451	AX356451	Sequence
C 117	20	1.3	1271	8	AY040022	C 190	20	1.3	10816	3	DME18278	Y18278	Drosophila
C 118	20	1.3	1337	8	AF274589	C 191	20	1.3	10855	6	AX346390	AX346390	Sequence
C 119	20	1.3	1351	9	BC007001	C 192	20	1.3	12359	6	AX346948	AX346948	Sequence
C 120	20	1.3	1401	3	AF338731	C 193	20	1.3	12733	6	AX344993	AX344993	Sequence
C 121	20	1.3	1440	9	BC001442	C 194	20	1.3	12927	2	AC107283	AC107283	Rattus no
C 122	20	1.3	1449	10	BC021947	C 195	20	1.3	14006	6	AX346860	AX346860	Sequence
C 123	20	1.3	1494	9	S7450453	C 196	20	1.3	15782	6	AX323521	AX323521	Sequence
C 124	20	1.3	1499	8	AY050825	C 197	20	1.3	15881	6	AX251764	AX251764	Sequence
C 125	20	1.3	1541	9	BC018788	C 198	20	1.3	15881	6	AX344178	AX344178	Sequence
C 126	20	1.3	1572	8	AF275752	C 199	20	1.3	15881	6	AX345162	AX345162	Sequence
C 127	20	1.3	1604	8	AF370210	C 200	20	1.3	15881	6	AX348575	AX348575	Sequence
C 128	20	1.3	1619	3	AY070934	C 201	20	1.3	16173	6	AX281358	AX281358	Sequence
C 129	20	1.3	1691	10	BC002192	C 202	20	1.3	17251	10	AB015623	AB015623	Mus muscu
C 130	20	1.3	1719	8	SPU16143	C 203	20	1.3	19946	9	AL583784	AL583784	Human DNA
C 131	20	1.3	1747	8	MCMDH1A	C 204	20	1.3	26203	2	PFMAL13P7	282187	Human DNA
C 132	20	1.3	1790	5	AF071570	C 205	20	1.3	27746	9	HS64F4	AC017410	Drosophila
C 133	20	1.3	1878	9	AK021503	C 206	20	1.3	29776	2	AC017410	AC017410	Drosophila
C 134	20	1.3	1897	9	AK054572	C 207	20	1.3	36030	8	F16L1	AC079942	Sequence
C 135	20	1.3	1926	9	AK025140	C 208	20	1.3	36034	8	AP001296	AP001296	Sequence
C 136	20	1.3	1943	3	DDV14	C 209	20	1.3	36252	9	AP001056	AP001056	Homo sapi
C 137	20	1.3	1957	9	HSB420507	C 210	20	1.3	36430	8	AC010506	AC010506	Sequence
C 138	20	1.3	1958	9	AK026674	C 211	20	1.3	37228	8	SPC757	AL031882	S. pombe
C 139	20	1.3	2113	10	AF418207	C 212	20	1.3	38237	2	AC013904	AC013904	Drosophila
C 140	20	1.3	2173	6	HSU28833	C 213	20	1.3	38478	2	AC019893	AC019893	Drosophila
C 141	20	1.3	2174	6	AK034241	C 214	20	1.3	41525	2	AC110343	AC110343	Rattus no
C 142	20	1.3	2227	9	HSU85266	C 215	20	1.3	43089	2	AC017875	AC017875	Drosophila
C 143	20	1.3	2284	9	HSU85265	C 216	20	1.3	43919	9	AL446023	AL446023	Human DNA
C 144	20	1.3	2346	3	AF243046	C 217	20	1.3	43983	9	AL136970	AL136970	Human DNA
C 145	20	1.3	2346	3	HSU85267	C 218	20	1.3	44623	9	AL138964	AL138964	Human DNA
C 146	20	1.3	2581	3	DDIGP138A	C 219	20	1.3	47090	9	AC092310	AC092310	Homo sapi

220	20	1.3	47090	9	AC093233	Homo sapi	c 293	20	1.3	110000	2	AL35978_1	Continuation (2 of
c 221	20	1.3	48282	2	AC024228	Arabidops	c 294	20	1.3	111051	9	AL353806	AL353806 Human DNA
c 222	20	1.3	49491	2	AC013847	Drosophil	c 295	20	1.3	112401	9	AL606462	AL606462 Human DNA
223	20	1.3	56631	9	HS702M17	AL033427	296	20	1.3	112573	9	AL353151	AL353151 Human DNA
224	20	1.3	59841	2	AC102369	Human DNA	297	20	1.3	113253	2	AC092356	AC092356 Human DNA
225	20	1.3	61159	2	AC021911	Homo sapi	c 298	20	1.3	113335	9	AL596268	AL596268 Human DNA
c 226	20	1.3	61278	2	AC024308	AC024308 Homo sapi	c 299	20	1.3	113515	6	AX347076	AX347076 Sequence
c 227	20	1.3	61514	2	AC087697	Homo sapi	300	20	1.3	114016	9	AL365396	AL365396 Human DNA
228	20	1.3	62497	2	AC103865	AC103865 Homo sapi	301	20	1.3	114604	9	AL391376	AL391376 Human DNA
229	20	1.3	63859	2	AC064794	Homo sapi	302	20	1.3	116215	9	AC004130	AC004130 Homo sapi
230	20	1.3	64737	2	AC104382	AC104382 Homo sapi	c 303	20	1.3	117613	9	AC092784	AC092784 Homo sapi
c 231	20	1.3	64737	2	AC104382	Homo sapi	304	20	1.3	121101	2	AC010931	AC010931 Homo sapi
c 232	20	1.3	66206	2	AC023324	AC023324 Homo sapi	305	20	1.3	123391	9	HS7902P15	HS7902P15 Human DNA
c 233	20	1.3	66259	9	AL390070	AL390070 Human DNA	c 306	20	1.3	125351	2	AC104642	AC104642 TYRpanoso
234	20	1.3	66352	2	AC110047	AC110047 Homo sapi	307	20	1.3	127370	9	AL451044	AL451044 Human DNA
235	20	1.3	67780	2	AC087275	AC087275 Homo sapi	c 308	20	1.3	127462	8	F5D14	AC007767 Sequence
c 236	20	1.3	68851	2	AC101183	AC101183 Mus muscu	c 309	20	1.3	128143	10	AC003063	AC003063 Mus muscu
c 237	20	1.3	68959	9	HSJ735P11	AL049783 Human DNA	310	20	1.3	128328	9	AC005839	AC005839 Homo sapi
238	20	1.3	68997	2	AC080153	AC080153 Homo sapi	c 311	20	1.3	128335	9	AC093864	AC093864 Homo sapi
c 239	20	1.3	70416	9	AC095039	AC095039 Homo sapi	c 312	20	1.3	128703	9	AF207550	AF207550 Homo sapi
c 240	20	1.3	70800	2	AC099950	AC099950 Mus muscu	313	20	1.3	129046	9	AC008835	AC008835 Homo sapi
c 241	20	1.3	71092	2	AL590732	AL590732 Human DNA	c 314	20	1.3	129129	9	AL513324	AL513324 Human DNA
242	20	1.3	71142	2	AC016305	AC016305 Homo sapi	c 315	20	1.3	132122	9	AC096582	AC096582 Homo sapi
c 243	20	1.3	73778	6	AX344562	AX344562 Sequence	c 316	20	1.3	132493	9	AC004583	AC004583 Human Chr
c 244	20	1.3	75609	9	AC006456	AC006456 Homo sapi	c 317	20	1.3	132809	10	AC087262	AC087262 Rattus no
245	20	1.3	76618	3	AC004657	AC004657 Drosophil	318	20	1.3	133368	9	AL627255	AL627255 Dantio rer
c 246	20	1.3	76808	2	AC014187	AL444198 Human DNA	319	20	1.3	134156	9	AL357083	AL357083 Human DNA
247	20	1.3	78404	9	AL445198	AL445198 Homo sapi	320	20	1.3	134210	9	AC005052	AC005052 Homo sapi
c 248	20	1.3	79432	2	AC014363	AC014363 Drosophil	c 321	20	1.3	134278	9	AP001822	AP001822 Homo sapi
c 249	20	1.3	82098	9	AC006252	AC006252 Homo sapi	c 322	20	1.3	134376	2	AC0073440	AC0073440 Homo sapi
c 250	20	1.3	83218	9	AL359980	AL359980 Human DNA	c 323	20	1.3	135053	2	AC094957	AC094957 Rattus no
251	20	1.3	83547	2	AC106849	AC106849 Homo sapi	c 324	20	1.3	135240	9	AL136362	AL136362 Human DNA
c 252	20	1.3	84388	9	AL589984	AL589984 Human DNA	325	20	1.3	136862	9	AC005007	AC005007 Homo sapi
c 253	20	1.3	86110	2	AC096145	AC096145 Rattus no	c 326	20	1.3	137184	9	AL353692	AL353692 Human DNA
254	20	1.3	87402	9	AP001426	AP001426 Homo sapi	c 327	20	1.3	137508	2	AC106622	AC106622 Rattus no
c 255	20	1.3	88118	2	AL645810	AL645810 Dantio rer	328	20	1.3	137515	2	DMBR12B32	DMBR12B32 Drosophil
c 256	20	1.3	89017	2	AC108387	AL121909 Human DNA	c 329	20	1.3	139773	2	AC067806	AC067806 Homo sapi
c 257	20	1.3	89251	9	HSJ812P3	AC094673 Rattus no	c 330	20	1.3	139773	2	CNS01DP0	CNS01DP0 Human chr
258	20	1.3	89369	2	AC094673	AC094673 Rattus no	331	20	1.3	140073	9	AL356415	AL356415 Human DNA
259	20	1.3	89684	9	AC064829	AC064829 Homo sapi	c 332	20	1.3	140999	9	AL356415	AL356415 Human DNA
260	20	1.3	90343	9	AC090287	AC090287 Homo sapi	c 333	20	1.3	141547	2	AC098660	AC098660 Rattus no
c 261	20	1.3	92558	9	HS341E18	AL031178 Human DNA	334	20	1.3	141794	2	AC093455	AC093455 Homo sapi
c 262	20	1.3	93443	9	HS063630	U63630 Homo sapien	c 335	20	1.3	143427	2	AC002432	AC002432 Human BAC
c 263	20	1.3	93805	9	HS424E5	AP000246 Homo sapi	c 336	20	1.3	143618	9	AL137856	AL137856 Human DNA
c 264	20	1.3	94212	9	AP000246	AL360083 Human DNA	c 337	20	1.3	145028	2	AC073875	AC073875 Homo sapi
c 265	20	1.3	94963	9	AL360083	AL360083 Rattus no	338	20	1.3	145189	2	AC021968	AC021968 Homo sapi
266	20	1.3	95911	2	AC094668	AC094668 Rattus no	339	20	1.3	145366	2	AC024945	AC024945 Homo sapi
c 267	20	1.3	96268	2	AC103047	AC103047 Rattus no	c 340	20	1.3	145380	9	AL357129	AL357129 Human DNA
c 268	20	1.3	96843	9	AC009496	AC009496 Homo sapi	c 341	20	1.3	146565	2	AL355542	AL355542 Homo sapi
c 269	20	1.3	99249	9	AL161916	AL161916 Human DNA	c 342	20	1.3	146814	2	AC022320	AC022320 Homo sapi
270	20	1.3	100000	9	AP000021	AP000021 Homo sapi	c 343	20	1.3	147465	2	AC093977	AC093977 Rattus no
c 271	20	1.3	100000	9	AP000070	AP000070 Homo sapi	344	20	1.3	148267	9	AC067940	AC067940 Homo sapi
c 272	20	1.3	100000	9	AP000130	AP000130 Homo sapi	345	20	1.3	148418	2	AC018258	AC018258 Drosophil
273	20	1.3	100000	9	AP000163	AP000163 Homo sapi	c 346	20	1.3	148777	5	AL590150	AL590150 Zebrafish
c 274	20	1.3	100000	9	AP000208	AP000208 Homo sapi	347	20	1.3	148847	2	AC023101	AC023101 Homo sapi
c 275	20	1.3	102343	2	AC106551	AC106551 Rattus no	348	20	1.3	149322	2	AC016039	AC016039 Homo sapi
c 276	20	1.3	104762	2	AC108086	AC108086 Homo sapi	349	20	1.3	149827	2	AL390880	AL390880 Homo sapi
c 277	20	1.3	105100	9	HS67A5	AL038594 Human DNA	c 350	20	1.3	149918	2	AC015504	AC015504 Homo sapi
c 278	20	1.3	105604	9	AC084013	AC084013 Homo sapi	351	20	1.3	149953	2	AC095571	AC095571 Rattus no
c 279	20	1.3	105797	9	AC011345	AC011345 Homo sapi	352	20	1.3	149995	9	AC016722	AC016722 Homo sapi
c 280	20	1.3	106341	9	AL358235	AL358235 Human DNA	353	20	1.3	150651	9	AL590559	AL590559 Human DNA
c 281	20	1.3	106506	9	AL358235	AL358235 Human DNA	c 354	20	1.3	150747	2	AC016230	AC016230 Homo sapi
c 282	20	1.3	106556	9	HSJ512E2	AL121787 Human DNA	355	20	1.3	152313	9	AL591594	AL591594 Human DNA
283	20	1.3	106676	2	AC011173	AC011173 Homo sapi	356	20	1.3	152555	9	AL354830	AL354830 Human DNA
c 284	20	1.3	106967	2	AC094530	AC094530 Rattus no	c 357	20	1.3	153190	9	AC068597	AC068597 Homo sapi
285	20	1.3	107567	2	AC011481	AC011481 Homo sapi	c 358	20	1.3	153190	9	AC026418	AC026418 Homo sapi
286	20	1.3	108967	9	AL353138	AL353138 Human DNA	c 359	20	1.3	153198	2	AL353327	AL353327 Homo sapi
287	20	1.3	109398	9	AC044817	AC044817 Homo sapi	c 360	20	1.3	153547	2	AC004926	AC004926 Homo sapi
288	20	1.3	109488	9	HS298J15	AL031774 Human DNA	361	20	1.3	153704	9	AC008567	AC008567 Homo sapi
289	20	1.3	110000	2	AC003656_1	Continuation (2 of	c 362	20	1.3	154036	9	AC008068	AC008068 Homo sapi
290	20	1.3	110000	2	AC073702_3	Continuation (4 of	363	20	1.3	154120	2	AP002472	AP002472 Homo sapi
291	20	1.3	110000	2	AC010536_1	Continuation (2 of	364	20	1.3	154452	2	AC106852	AC106852 Homo sapi
c 292	20	1.3	110000	2	AL355364_0	AL355364 Homo sapi	365	20	1.3	154679	2	AC022019	AC022019 Homo sapi

366	20	1.3	155124	2	AL356112	AL356112 Homo sapi	c 439	20	1.3	169159	2	AC023035	AC023035 Homo sapi
367	20	1.3	155157	9	AL360013	AL360013 Human DNA	440	20	1.3	169494	2	AC010299	AC010299 Homo sapi
368	20	1.3	155584	9	CNS07ERO	AL513143 Homo chr	441	20	1.3	169546	2	AC004157	AC004157 Plasmodiu
369	20	1.3	155666	9	AC008696	Homo sapi	442	20	1.3	169890	2	AL662872	AL662872 Homo sapi
370	20	1.3	155767	9	AP001603	Homo sapi	443	20	1.3	170071	2	AC094342	AC094342 Rattus no
371	20	1.3	155845	2	AC092122	Homo sapi	c 444	20	1.3	170105	3	AC007756	AC007756 Drosophil
372	20	1.3	155867	2	HS404H4	AL0031661 Human DNA	c 445	20	1.3	170163	3	AC009457	AC009457 Drosophil
373	20	1.3	156060	2	AC004153	Plasmodiu	446	20	1.3	170249	9	AC068799	AC068799 Homo sapi
374	20	1.3	156499	2	HS279F22	AL662879 Homo sapi	447	20	1.3	170270	2	AP001548	AP001548 Homo sapi
375	20	1.3	157015	2	AP001365	Homo sapi	448	20	1.3	170392	2	AC021523	AC021523 Homo sapi
376	20	1.3	157228	2	AC091495	Homo sapi	c 449	20	1.3	170413	2	AC090405	AC090405 Homo sapi
377	20	1.3	157336	2	AC079787	Homo sapi	450	20	1.3	170513	2	AC090213	AC090213 Homo sapi
378	20	1.3	157394	2	AL138790	Homo sapi	451	20	1.3	170740	9	HSB80K6	AL121902 Human DNA
379	20	1.3	157439	2	AC026355	Homo sapi	452	20	1.3	170760	2	AC004071	AC004071 Homo sapi
380	20	1.3	157454	2	AC037458	Homo sapi	c 453	20	1.3	171016	2	AC018357	AC018357 Homo sapi
381	20	1.3	157485	2	AC079056	Homo sapi	c 454	20	1.3	171069	2	AC109559	AC109559 Homo sapi
382	20	1.3	157508	9	AC018422	Homo sapi	c 455	20	1.3	171386	2	AP001198	AP001198 Homo sapi
383	20	1.3	157711	9	AC055738	Homo sapi	c 456	20	1.3	171790	2	AC011106	AC011106 Homo sapi
384	20	1.3	157848	9	AC079763	Homo sapi	c 457	20	1.3	172112	2	AC060822	AC060822 Homo sapi
385	20	1.3	158143	9	AL157384	Human DNA	458	20	1.3	172138	2	AC064796	AC064796 Homo sapi
386	20	1.3	158739	2	AC027402	Homo sapi	c 459	20	1.3	172161	8	AP003301	AP003301 Oryza sat
387	20	1.3	158886	2	AC016944	Homo sapi	c 460	20	1.3	172437	2	AC096437	AC096437 Rattus no
388	20	1.3	159287	9	AP003694	Homo sapi	461	20	1.3	172759	9	AC025589	AC025589 Homo sapi
389	20	1.3	159709	2	AC078806	Homo sapi	462	20	1.3	173211	9	AL353732	AL353732 Human DNA
390	20	1.3	160114	2	AL157780	Human DNA	463	20	1.3	173292	9	CNS01DVA	AL133502 Human chr
391	20	1.3	160391	2	AC010786	Homo sapi	c 464	20	1.3	173913	2	AC079628	AC079628 Homo sapi
392	20	1.3	160525	2	AC027074	Homo sapi	c 465	20	1.3	174232	2	AC026251	AC026251 Homo sapi
393	20	1.3	160536	9	AL355984	Human DNA	c 466	20	1.3	174576	2	AC105180	AC105180 Homo sapi
394	20	1.3	160559	2	AC022807	Homo sapi	467	20	1.3	174640	2	AC024479	AC024479 Homo sapi
395	20	1.3	161133	2	AC019295	Homo sapi	c 468	20	1.3	174645	2	AC091111	AC091111 Homo sapi
396	20	1.3	161164	2	AC022278	Homo sapi	c 469	20	1.3	174657	2	AC1108479	AC1108479 Drosophil
397	20	1.3	161298	2	AP002858	Homo sapi	c 470	20	1.3	174902	3	AC110071	AC110071 Homo sapi
398	20	1.3	161359	2	AC090025	Homo sapi	471	20	1.3	175218	2	AC027031	AC027031 Homo sapi
399	20	1.3	161624	2	AP001493	Homo sapi	472	20	1.3	175353	3	AC023708	AC023708 Drosophil
400	20	1.3	161778	2	AC097003	Rattus no	473	20	1.3	175365	2	AC069435	AC069435 Homo sapi
401	20	1.3	162064	9	AL390882	Human DNA	474	20	1.3	175515	2	AC098128	AC098128 Rattus no
402	20	1.3	162260	2	AC021236	Homo sapi	c 475	20	1.3	175770	9	AC026523	AC026523 Homo sapi
403	20	1.3	162539	2	AL390720	Homo sapi	c 476	20	1.3	175850	2	AP001399	AP001399 Homo sapi
404	20	1.3	162752	9	AC099559	Homo sapi	c 477	20	1.3	175916	2	AC007747	AC007747 Homo sapi
405	20	1.3	162837	2	AL451137	Human DNA	478	20	1.3	176161	2	AC021173	AC021173 Homo sapi
406	20	1.3	163314	9	AL590557	Homo sapi	c 479	20	1.3	176300	2	AC026311	AC026311 Homo sapi
407	20	1.3	163386	9	AC021659	Homo sapi	480	20	1.3	176528	2	AC074187	AC074187 Homo sapi
408	20	1.3	163397	2	CNS01RG2	Homo sapi	481	20	1.3	176668	2	AC027096	AC027096 Homo sapi
409	20	1.3	163537	2	AC080014	Homo sapi	c 482	20	1.3	176734	2	AC094887	AC094887 Rattus no
410	20	1.3	163690	2	AC083781	Homo sapi	c 483	20	1.3	176783	9	AL139281	AL139281 Human DNA
411	20	1.3	163954	9	AC090959	Homo sapi	484	20	1.3	177037	9	AL163542	AL163542 Human DNA
412	20	1.3	164173	2	AC016072	Homo sapi	c 485	20	1.3	177338	2	AC095133	AC095133 Rattus no
413	20	1.3	164452	9	CNS01RGP	Human chr	486	20	1.3	177444	2	AC084238	AC084238 Mus muscu
414	20	1.3	164507	2	AC012067	Homo sapi	487	20	1.3	177464	9	AP002853	AP002853 Homo sapi
415	20	1.3	165556	2	AC087190	Homo sapi	488	20	1.3	177536	9	AC010686	AC010686 Homo sapi
416	20	1.3	165873	9	AC066593	Homo sapi	c 489	20	1.3	177646	2	AC026317	AC026317 Homo sapi
417	20	1.3	165924	2	AC107024	Homo sapi	c 490	20	1.3	177777	9	AC007342	AC007342 Homo sapi
418	20	1.3	166399	2	AC012584	Homo sapi	c 491	20	1.3	177984	9	AC015700	AC015700 Homo sapi
419	20	1.3	166892	9	AC007463	Homo sapi	c 492	20	1.3	178053	2	AC021112	AC021112 Homo sapi
420	20	1.3	166952	2	AC073856	Homo sapi	c 493	20	1.3	178201	2	AC025888	AC025888 Homo sapi
421	20	1.3	167014	2	AC097705	Rattus no	494	20	1.3	178203	9	AC008780	AC008780 Homo sapi
422	20	1.3	167478	2	AC011148	Homo sapi	495	20	1.3	178526	2	AC106037	AC106037 Homo sapi
423	20	1.3	167512	2	AC015708	Homo sapi	c 496	20	1.3	178637	2	AC099337	AC099337 Homo sapi
424	20	1.3	167568	2	AC063975	Homo sapi	c 497	20	1.3	178717	2	AC024396	AC024396 Homo sapi
425	20	1.3	167589	2	AC023235	Homo sapi	c 498	20	1.3	178756	2	AC1044420	AC1044420 Homo sapi
426	20	1.3	167642	9	CNS07ECT	Human chr	499	20	1.3	179137	2	AC016641	AC016641 Homo sapi
427	20	1.3	167910	2	AC023203	Homo sapi	c 500	20	1.3	179141	2	AC019091	AC019091 Homo sapi
428	20	1.3	167962	2	AC073484	Homo sapi	c 501	20	1.3	179420	2	AL592296	AL592296 Homo sapi
429	20	1.3	167996	9	AC021753	Homo sapi	c 502	20	1.3	179698	2	AC021848	AC021848 Homo sapi
430	20	1.3	168193	2	CNS05TE2	Human chr	c 503	20	1.3	179774	2	AC026223	AC026223 Homo sapi
431	20	1.3	168208	2	AL646091	Mus muscu	504	20	1.3	179892	3	AC023691	AC023691 Drosophil
432	20	1.3	168265	2	AC105919	Homo sapi	505	20	1.3	179929	2	AC006270	AC006270 Homo sapi
433	20	1.3	168306	2	AL627308	Homo sapi	506	20	1.3	180034	2	AC072053	AC072053 Homo sapi
434	20	1.3	168571	3	AC104510	Drosophil	507	20	1.3	180109	2	AC026896	AC026896 Homo sapi
435	20	1.3	168703	2	AC091759	Sus scrofa	c 508	20	1.3	180533	2	AL589943	AL589943 Homo sapi
436	20	1.3	168768	2	AL672047	Mus muscu	509	20	1.3	180673	10	AC068627	AC068627 Mus muscu
437	20	1.3	168818	2	AC013496	Homo sapi	c 510	20	1.3	180778	9	AL355596	AL355596 Human DNA
438	20	1.3	169114	2	AC096867	Rattus no	c 511	20	1.3	180849	30	AC023996	AC023996 Homo sapi

512	20	1.3	181148	2	AC018413	Homo sapi	585	20	1.3	199841	2	AC099729	Mus muscu
513	20	1.3	181252	9	AC010776	Homo sapi	586	20	1.3	199892	2	AC020997	AC020997
514	20	1.3	181438	2	AC023047	Homo sapi	587	20	1.3	199988	2	AC023545	AC023545
515	20	1.3	181563	2	AL670035	Mus muscu	588	20	1.3	200780	2	AC098827	AC098827
516	20	1.3	181716	2	AC068285	Homo sapi	589	20	1.3	200789	2	AL672090	AL672090
517	20	1.3	181823	2	AC040167	Homo sapi	590	20	1.3	201408	9	AC090644	AC090644
518	20	1.3	181935	2	AC079122	AC079122	591	20	1.3	201451	2	AC079623	AC079623
519	20	1.3	181999	2	AC106519	Rattus no	592	20	1.3	201579	2	AC068773	AC068773
520	20	1.3	182048	2	AC093293	Homo sapi	593	20	1.3	201757	10	AC061963	AC061963
521	20	1.3	182102	2	AC097274	Mus muscu	594	20	1.3	202423	2	AL604025	AL604025
522	20	1.3	182103	2	AC016998	Homo sapi	595	20	1.3	202550	2	AL391832	AL391832
523	20	1.3	182321	2	AP001357	Homo sapi	596	20	1.3	202638	2	AC109825	AC109825
524	20	1.3	182588	2	AC069514	Homo sapi	597	20	1.3	203523	2	AC068127	AC068127
525	20	1.3	182834	2	AC027523	Homo sapi	598	20	1.3	203842	9	AC010789	AC010789
526	20	1.3	183424	2	AC093387	Homo sapi	599	20	1.3	203853	2	AC098649	AC098649
527	20	1.3	183439	3	AC007809	Drosophi	600	20	1.3	203946	2	AC069445	AC069445
528	20	1.3	184043	9	AC097505	Homo sapi	601	20	1.3	204153	9	AC074091	AC074091
529	20	1.3	184092	2	AC018715	Homo sapi	602	20	1.3	204340	9	AC091103	AC091103
530	20	1.3	184454	3	AC104626	Drosophi	603	20	1.3	204767	2	AC099781	AC099781
531	20	1.3	184455	9	AC010740	Homo sapi	604	20	1.3	205815	2	AC008759	AC008759
532	20	1.3	184517	2	AC091758	Sus scrof	605	20	1.3	205831	2	AC074192	AC074192
533	20	1.3	184568	2	AC009558	Homo sapi	606	20	1.3	205932	2	AC073788	AC073788
534	20	1.3	184643	2	AC087650	Homo sapi	607	20	1.3	205993	2	AC011952	AC011952
535	20	1.3	184663	2	AC012113	Homo sapi	608	20	1.3	206743	2	AL663081	AL663081
536	20	1.3	184689	9	AL442125	Human DNA	609	20	1.3	207439	2	AC110615	AC110615
537	20	1.3	184916	2	AC068998	Mus muscu	610	20	1.3	207897	2	AL590389	AL590389
538	20	1.3	185024	2	AC027649	Mus muscu	611	20	1.3	207945	9	CNS01D5S	CNS01D5S
539	20	1.3	185414	2	AL355526	Homo sapi	612	20	1.3	209512	9	CNS01BOV	CNS01BOV
540	20	1.3	185602	2	AC097585	Sus scrof	613	20	1.3	210107	9	AL365214	AL365214
541	20	1.3	186047	2	AC074257	Homo sapi	614	20	1.3	211627	9	AL161651	AL161651
542	20	1.3	186150	2	AP001544	Homo sapi	615	20	1.3	212506	2	AC079516	AC079516
543	20	1.3	186166	2	AL391279	Homo sapi	616	20	1.3	213204	2	AC078895	AC078895
544	20	1.3	186656	2	AL592143	Homo sapi	617	20	1.3	213901	2	AC093472	AC093472
545	20	1.3	186716	2	AC025863	AC025863	618	20	1.3	214567	2	AC084316	AC084316
546	20	1.3	186716	3	AC023688	Drosophi	619	20	1.3	214656	2	AF235101	AF235101
547	20	1.3	187704	2	AC025586	Mus muscu	620	20	1.3	214875	10	AC015584	AC015584
548	20	1.3	187806	2	AC076971	Homo sapi	621	20	1.3	215196	2	AL603705	AL603705
549	20	1.3	187852	2	AC107293	Homo sapi	622	20	1.3	215653	2	AC073742	AC073742
550	20	1.3	188582	2	AC013720	Homo sapi	623	20	1.3	216200	9	AP000764	AP000764
551	20	1.3	188863	9	AC007345	Homo sapi	624	20	1.3	216766	2	AC103953	AC103953
552	20	1.3	189134	9	H0K0004682	AC004682	625	20	1.3	216988	10	AC084217	AC084217
553	20	1.3	189427	9	AC023286	Mus muscu	626	20	1.3	217421	9	AC006211	AC006211
554	20	1.3	189801	30	AC027447	Homo sapi	627	20	1.3	217730	2	AC087440	AC087440
555	20	1.3	189893	9	AP000999	Homo sapi	628	20	1.3	217747	2	AC105448	AC105448
556	20	1.3	189903	2	AC012542	Homo sapi	629	20	1.3	218589	2	AC025083	AC025083
557	20	1.3	190856	2	AC009160	Homo sapi	630	20	1.3	219175	9	AC092641	AC092641
558	20	1.3	190856	2	AC024505	Homo sapi	631	20	1.3	219720	9	AC023364	AC023364
559	20	1.3	191464	9	CNS01DPI	Human chr	632	20	1.3	221409	3	AE003730	AE003730
560	20	1.3	191485	2	AC091467	Mus muscu	633	20	1.3	221688	2	AE010323	AE010323
561	20	1.3	191655	9	AC020661	Homo sapi	634	20	1.3	221830	3	AC009393	AC009393
562	20	1.3	191911	9	AC098965	Homo sapi	635	20	1.3	222329	10	AC074310	AC074310
563	20	1.3	191911	9	AC098965	Homo sapi	636	20	1.3	224010	2	AP001848	AP001848
564	20	1.3	192179	2	AC097365	Mus muscu	637	20	1.3	226199	3	AC084069	AC084069
565	20	1.3	192189	2	AC107212	Homo sapi	638	20	1.3	226323	3	AE003697	AE003697
566	20	1.3	192358	3	AC099028	Drosophi	639	20	1.3	226825	2	AC073912	AC073912
567	20	1.3	192581	2	PFMAL13P1	AL049180	640	20	1.3	227348	2	AC092851	AC092851
568	20	1.3	192819	2	AC012280	Plasmodiu	641	20	1.3	228156	9	AC022150	AC022150
569	20	1.3	193098	9	AL137140	Homo sapi	642	20	1.3	229655	3	AE003707	AE003707
570	20	1.3	193117	2	AC106721	Human DNA	643	20	1.3	231661	2	AC073770	AC073770
571	20	1.3	194041	2	AL662912	Mus muscu	644	20	1.3	232744	3	AE003705	AE003705
572	20	1.3	194198	2	AC011774	Homo sapi	645	20	1.3	234235	2	AC099303	AC099303
573	20	1.3	194418	9	AF235098	Homo sapi	646	20	1.3	235150	2	AC073776	AC073776
574	20	1.3	194487	2	AP002896	Homo sapi	647	20	1.3	237119	2	AC017740	AC017740
575	20	1.3	195102	9	AL353795	Human DNA	648	20	1.3	239108	2	AL606661	AL606661
576	20	1.3	195250	2	AC023607	Mus muscu	649	20	1.3	240294	2	AC044892	AC044892
577	20	1.3	195250	2	AC023607	Mus muscu	650	20	1.3	241857	2	AL645911	AL645911
578	20	1.3	195743	2	AC016774	Homo sapi	651	20	1.3	242965	2	AC023175	AC023175
579	20	1.3	196349	2	AC019112	Homo sapi	652	20	1.3	247337	2	AC099070	AC099070
580	20	1.3	196413	9	AC097269	Pen trogl	653	20	1.3	258174	2	AC079429	AC079429
581	20	1.3	196828	2	AC007363	Homo sapi	654	20	1.3	259978	2	AC073819	AC073819
582	20	1.3	198829	2	AL391668	Homo sapi	655	20	1.3	261090	3	AE003798	AE003798
583	20	1.3	199377	2	AC084294	Mus muscu	656	20	1.3	265815	2	AC027308	AC027308
584	20	1.3	199730	2	AC083940	Homo sapi	657	20	1.3	279090	2	AC091290	AC091290





804	19	1.2	1008	8	BNPBP4	X90730 B.napus mRN	877	19	1.2	1403	8	ATU83624	U83624 Arabidopsis
805	19	1.2	1009	3	AF030628	AF030628 Plasmodiu	878	19	1.2	1414	8	ATH7587	AF007587 Arabidops
806	19	1.2	1011	3	AY075365	AY075365 Drosophi	879	19	1.2	1421	10	AF051425	AF051425 Rattus no
807	19	1.2	1021	9	BC005200	BC005200 Homo sapi	880	19	1.2	1439	8	AY065232	AY065232 Arabidops
808	19	1.2	1035	9	BC002828	BC002828 Homo sapi	881	19	1.2	1439	3	AY058384	AY058384 Drosophi
809	19	1.2	1036	8	BVCHLORP	Y13865 Beta vulgar	882	19	1.2	1457	9	BC003017	BC003017 Homo sapi
810	19	1.2	1041	3	AF231352	AF231352 Ctenoceph	883	19	1.2	1458	9	AF070645	AF070645 Homo sapi
811	19	1.2	1045	8	AY034967	AY034967 Arabidops	884	19	1.2	1466	9	BC000460	BC000460 Homo sapi
812	19	1.2	1047	8	BC005132	BC005132 Homo sapi	885	19	1.2	1477	8	AF028842	AF028842 Arabidops
813	19	1.2	1048	8	AF297521	AF297521 Prunus av	886	19	1.2	1481	3	AY058287	AY058287 Drosophi
814	19	1.2	1054	10	BC021770	BC021770 Mus muscu	887	19	1.2	1482	5	AB071891	AB071891 Letchentr
815	19	1.2	1065	3	AY061524	AY061524 Drosophi	888	19	1.2	1507	9	AF090988	AF090988 Homo sapi
816	19	1.2	1080	8	NTT232AF	X62427 N.tabacum m	889	19	1.2	1512	8	AB064519	AB064519 Oryza sat
817	19	1.2	1085	5	AF270790	AF270790 Orosachrom	890	19	1.2	1516	9	BC020631	BC020631 Homo sapi
818	19	1.2	1093	8	AB047397	AB047397 Zygosaech	891	19	1.2	1531	9	BC001494	BC001494 Homo sapi
819	19	1.2	1096	9	BC004277	BC004277 Homo sapi	892	19	1.2	1541	8	AF370323	AF370323 Arabidops
820	19	1.2	1102	9	BC012033	BC012033 Homo sapi	893	19	1.2	1542	6	AF370332	AF370332 Arabidops
821	19	1.2	1107	9	HSMB02425	AL157420 Homo sapi	894	19	1.2	1546	5	AF102890	AF102890 Gadus mor
822	19	1.2	1110	5	SSAJ4825	AF004825 Salmo sal	895	19	1.2	1556	9	BC008009	BC008009 Homo sapi
823	19	1.2	1113	8	AF043519	AF043519 Arabidops	896	19	1.2	1566	9	BC004983	BC004983 Homo sapi
824	19	1.2	1114	9	BC015039	BC015039 Homo sapi	897	19	1.2	1569	9	AK026099	AK026099 Homo sapi
825	19	1.2	1115	3	AY060811	AY060811 Drosophi	898	19	1.2	1570	3	AY069282	AY069282 Drosophi
826	19	1.2	1137	10	BC002003	BC002003 Mus muscu	899	19	1.2	1578	6	AK138898	AK138898 Sequence
827	19	1.2	1140	3	AY069465	AY069465 Drosophi	900	19	1.2	1578	6	E29750	E29750 cDNA clone
828	19	1.2	1150	5	AF140281	AF140281 Cirrhinus	901	19	1.2	1582	3	AY052107	AY052107 Drosophi
829	19	1.2	1160	10	AF408432	AF408432 Mus muscu	902	19	1.2	1584	9	BC002601	BC002601 Homo sapi
830	19	1.2	1161	10	BC002332	BC002332 Mus muscu	903	19	1.2	1584	10	BC016573	BC016573 Mus muscu
831	19	1.2	1163	5	SSAJ4824	AI004824 Salmo sal	904	19	1.2	1598	5	XLXMT8	X57224 X.laevis xw
832	19	1.2	1168	9	AF165522	AF165522 Homo sapi	905	19	1.2	1631	9	BC012134	BC012134 Homo sapi
833	19	1.2	1171	3	AY075439	AY075439 Drosophi	906	19	1.2	1642	8	AY062178	AY062178 Oryza sat
834	19	1.2	1172	8	AY072075	AY072075 Arabidops	907	19	1.2	1644	8	AY072108	AY072108 Arabidops
835	19	1.2	1177	8	AY063837	AY063837 Arabidops	908	19	1.2	1648	8	AY050893	AY050893 Arabidops
836	19	1.2	1182	8	AF451952	AF451952 Arabidops	909	19	1.2	1654	8	AY056117	AY056117 Arabidops
837	19	1.2	1186	9	BC000793	BC000793 Homo sapi	910	19	1.2	1656	5	AB010713	AB010713 Cirrhinus
838	19	1.2	1196	4	CA035049	U33049 Ovis aries	911	19	1.2	1659	9	BC015312	BC015312 Homo sapi
839	19	1.2	1201	9	BC012589	BC012589 Homo sapi	912	19	1.2	1665	9	AK024914	AK024914 Homo sapi
840	19	1.2	1202	9	HUMETFC4	LI8960 Human prote	913	19	1.2	1667	14	AB021977	AB021977 Rinderpes
841	19	1.2	1203	8	AY063792	BC063792 Arabidops	914	19	1.2	1684	8	AF370136	AF370136 Arabidops
842	19	1.2	1204	10	BC005743	BC005743 Mus muscu	915	19	1.2	1687	3	AY052117	AY052117 Drosophi
843	19	1.2	1205	9	BC005840	BC005840 Homo sapi	916	19	1.2	1692	9	BC010359	BC010359 Homo sapi
844	19	1.2	1215	3	AY071536	AY071536 Drosophi	917	19	1.2	1700	8	AF166352	AF166352 Arabidops
845	19	1.2	1215	8	LPE243504	AI243504 Lolium pe	918	19	1.2	1701	9	HS421991	HS421991 Homo sapi
846	19	1.2	1223	3	DDIDISIA	J01282 S.lime mold	919	19	1.2	1701	9	HS421991	X44709 H.sapiens m
847	19	1.2	1223	8	GGN132215	AI132215 Gnetum gn	920	19	1.2	1703	9	AB048936	AB048936 Macaca fa
848	19	1.2	1224	8	AY050929	AY050929 Arabidops	921	19	1.2	1704	9	H009384S20	AF219908 Homo sapi
849	19	1.2	1226	9	AK026455	AK026455 Homo sapi	922	19	1.2	1708	9	AF070559	AF070559 Homo sapi
850	19	1.2	1239	10	MMCENIRA	X17373 M.musculus	923	19	1.2	1717	10	BC023168	BC023168 Mus muscu
851	19	1.2	1242	10	AF143369	AF143369 Mus muscu	924	19	1.2	1724	9	BC001779	BC001779 Homo sapi
852	19	1.2	1245	9	AF282025	AF282025 Homo sapi	925	19	1.2	1733	9	HSOCr6TR	Z18284 Homo sapien
853	19	1.2	1249	9	BC021697	BC021697 Homo sapi	926	19	1.2	1735	9	BC017836	BC017836 Homo sapi
854	19	1.2	1250	8	AY048692	AY048692 Arabidops	927	19	1.2	1739	9	AB048980	AB048980 Macaca fa
855	19	1.2	1261	9	BC021105	BC021105 Homo sapi	928	19	1.2	1742	10	BC016107	BC016107 Mus muscu
856	19	1.2	1266	9	HSMB02163	AL137452 Homo sapi	929	19	1.2	1744	3	AK024845	AK024845 Homo sapi
857	19	1.2	1270	9	BC019326	BC019326 Homo sapi	930	19	1.2	1753	3	AY061291	AY061291 Drosophi
858	19	1.2	1273	9	BC021667	BC021667 Homo sapi	931	19	1.2	1755	9	HSAF000974	AF000974 Human zyx
859	19	1.2	1286	10	AF061262	AF061262 Mus muscu	932	19	1.2	1777	9	HSTECTXT11	U43777 Human tyros
860	19	1.2	1300	9	AF005214	AF005214 Homo sapi	933	19	1.2	1806	10	MMU295253	AJ295253 Mus muscu
861	19	1.2	1302	9	BC008906	BC008906 Homo sapi	934	19	1.2	1816	9	AF273770	AF273770 Ctenorhab
862	19	1.2	1318	6	AX008671	AX008671 Sequence	935	19	1.2	1817	5	XLA011652	AI011652 Xenopus l
863	19	1.2	1318	8	PC0132438	AI132438 Phasaelus	936	19	1.2	1819	9	AY007790	AY007790 Homo sapi
864	19	1.2	1329	8	AY045905	AY045905 Arabidops	937	19	1.2	1827	8	CRO7332	AF007332 Catharant
865	19	1.2	1331	10	BC013440	BC013440 Mus muscu	938	19	1.2	1839	9	AK000096	AK000096 Homo sapi
866	19	1.2	1332	9	HSAA30565	AJ240565 Homo sapi	939	19	1.2	1841	3	AY060776	AY060776 Drosophi
867	19	1.2	1341	3	DME278068	AJ278068 Drosophi	940	19	1.2	1843	8	NTA271750	AJ271750 Nicotiana
868	19	1.2	1347	3	BC002596	BC002596 Homo sapi	941	19	1.2	1844	9	BC010741	BC010741 Homo sapi
869	19	1.2	1350	3	AY071533	AY071533 Drosophi	942	19	1.2	1858	9	AB046649	AB046649 Macaca fa
870	19	1.2	1362	8	AY063851	AY063851 Arabidops	943	19	1.2	1860	6	AF015566	AF015566 Dictyoste
871	19	1.2	1364	8	AK000039	AK000039 Homo sapi	944	19	1.2	1874	3	AY071374	AY071374 Drosophi
872	19	1.2	1367	8	AY065422	AY065422 Arabidops	945	19	1.2	1879	3	AY061829	AY061829 Drosophi
873	19	1.2	1371	8	AY035016	AY035016 Arabidops	946	19	1.2	1883	3	AY058766	AY058766 Drosophi
874	19	1.2	1385	3	AY058467	AY058467 Drosophi	947	19	1.2	1889	5	AB025103	AB025103 Gallus ga
875	19	1.2	1386	9	HSU77327	U77327 Human Ki-1/	948	19	1.2	1899	10	BC007468	BC007468 Mus muscu
876	19	1.2	1401	9	BC000932	BC000932 Homo sapi	949	19	1.2	1889	10	BC007468	BC007468 Mus muscu

c	950	19	1.2	1890	3	AY061032	AY061032 Drosophila
c	951	19	1.2	1891	8	AY056274	AY056274 Arabidops
c	952	19	1.2	1893	10	BC002211	BC002211 Mus muscu
c	953	19	1.2	1901	3	AY061150	AY061150 Drosophila
c	954	19	1.2	1908	9	BC022343	BC022343 Homo sapi
c	955	19	1.2	1917	3	PFA38237	AJ138237 Plasmodiu
c	956	19	1.2	1917	9	BC012605	BC012605 Homo sapi
c	957	19	1.2	1923	9	BC004159	BC004159 Homo sapi
c	958	19	1.2	1925	8	AF360263	AF360263 Arabidops
c	959	19	1.2	1927	3	AF041039	AF041039 Aplysia c
c	960	19	1.2	1932	9	BC005247	BC005247 Homo sapi
c	961	19	1.2	1933	6	E01979	E01979 DNA encodin
c	962	19	1.2	1936	8	AB059557	AB059557 Avena sat
c	963	19	1.2	1937	10	AB042807	AB042807 Mus muscu
c	964	19	1.2	1940	9	AB063000	AB063000 Macaca fa
c	965	19	1.2	1944	10	AF177756	AF177756 Mus muscu
c	966	19	1.2	1946	9	BC015829	BC015829 Homo sapi
c	967	19	1.2	1950	10	BC022903	BC022903 Mus muscu
c	968	19	1.2	1959	9	HSM802478	HSM802478 Homo sapi
c	969	19	1.2	1959	10	AF177755	AF177755 Mus muscu
c	970	19	1.2	1966	8	AY062625	AY062625 Arabidops
c	971	19	1.2	1971	9	AF068179	AF068179 Homo sapi
c	972	19	1.2	1972	8	AF355056	AF355056 Linum usi
c	973	19	1.2	1982	9	AK024775	AK024775 Homo sapi
c	974	19	1.2	1988	9	AY035400	AY035400 Homo sapi
c	975	19	1.2	1997	3	AF132144	AF132144 Drosophila
c	976	19	1.2	2007	3	ABR132102	AJ132102 Astropsect
c	977	19	1.2	2007	9	AF217985	AF217985 Homo sapi
c	978	19	1.2	2016	3	AY060433	AY060433 Drosophila
c	979	19	1.2	2021	8	AB029334	AB029334 Halocynth
c	980	19	1.2	2028	8	ZMTUG1	X83695 Z.mays mRNA
c	981	19	1.2	2046	14	AF404345	AF404345 Infectiou
c	982	19	1.2	2060	6	ABR171282	ABR171282 Sequence
c	983	19	1.2	2061	5	AB001573	AB001573 Oryzias l
c	984	19	1.2	2070	10	AF152996	AF152996 Mus muscu
c	985	19	1.2	2070	10	BC023105	BC023105 Mus muscu
c	986	19	1.2	2077	9	AB070001	AB070001 Macaca fa
c	987	19	1.2	2081	10	BC003989	BC003989 Mus muscu
c	988	19	1.2	2088	5	XLA7835	AJ007835 Xenopus L
c	989	19	1.2	2091	8	AY070036	AY070036 Arabidops
c	990	19	1.2	2096	8	AY035103	AY035103 Arabidops
c	991	19	1.2	2107	8	AF289952	X00798 Nicotiana d
c	992	19	1.2	2108	10	AF289953	AF289953 Mus muscu
c	993	19	1.2	2125	3	AY075207	AY075207 Drosophila
c	994	19	1.2	2128	3	AY060791	AY060791 Drosophila
c	995	19	1.2	2133	9	AF163577	AF163577 Homo sapi
c	996	19	1.2	2134	9	HSM802293	AL137554 Homo sapi
c	997	19	1.2	2138	10	BC011081	BC011081 Mus muscu
c	998	19	1.2	2146	8	AF224707	AF224707 Arabidops
c	999	19	1.2	2149	9	AB055285	AB055285 Macaca fa
c	1000	19	1.2	2161	9	AB062998	AB062998 Macaca fa

ALIGNMENTS

RESULT 1

AX113259

LOCUS AX113259

DEFINITION Sequence 1 from Patent EP1106689.

ACCESSION AX113259

VERSION AX113259.1

KEYWORDS GI:13939667

SOURCE fruit fly.

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1539)

Adamczewski,M., Methfessel,C. and Schulte,T.

Authors Nucleic acids encoding acetylcholine beta subunits from insects

Title Patent: EP 1106689-A 1 13-JUN-2001;

Journal BAYER AG (DE)

1539 bp

DNA

FEATURES	source	Location/Qualifiers
		1..1539
		/organism="Drosophila melanogaster"
		/db_xref="taxon:722"
CDS		43..1368
		/note="unamed protein product"
		/codon_start=1
		/protein_id="CAC37851.1"
		/db_xref="GI:13939668"
		/translation="MTTTPKIKAPVSGPGLLIQLMGLMLGITSVPGATFADPP NANVKADRLRHAGLFPTNDSVPVFDQTPNVSLEMYVYIDIDELNGKLTTCWML LRMDEERKVMQSQDYDNTQITLQKSEWYQIITLFDNDEGLAEVOTLSDHGF MMPAVYAYVCELEMLMMPHDKQSKLIKISGMGKVVLPENGTARGESLDDDLVQSF EMELVDSHAHVESODDYGYMEYETLTAQRSMGAVVYTPASCIYIALISAFMLP GKEKIMINGLLIIVIAAFVFAQLITPLSNPTLVYTESSTILYSVSTIVYIVYI YLAGCKHRRRLPEALRKLHLHSLTSMWLLSNPTLVYTESSTILYSVSTIVYIVYI SSPLGINTEYVPAKANQFDWALATVADRISFVSVSLAFILITRCSV"
BASE COUNT	352 a 476 c 411 g 300 t	
ORIGIN		
Query Match	100.0%	Score 1539; DB 6; Length 1539;
Best Local Similarity	100.0%	Pred. No. 0;
Matches 1539; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
QY	1	atccgcacgaaggtacacccgaaacaaaggcgcgctgaacaatgacagacatccccaag 60
Db	1	ATTGCGCAGAGAGGTACATCCGAAMCAAAAGCGCGCTGACATGACGACGACTCCAAAG 60
OY	61	ataaaggacccaagtttcggtgtccgtgtcgtgactgtccactactgtgtcgaatgtatgggagt 120
Db	61	ATPAAAGCACCAAGTTTCGGGTCTTGACTGCGCACATACCTGCTCAATGCTTAATGGGGATG 120
OY	121	ctctctatgagctactctccgtgcgcgaagcgccaatgccaacgcggagacccaagaagcgc 180
Db	121	CTTCTTATGGGGCTACTCTCCGTGCGCAGGCGGCATGCGCACCGCGGACCCCAAGAACGCC 180
OY	181	aatgtcaaggcccttgatctgctccacgcgcgcgtgttcaogaactacgaacgaagatgt 240
Db	181	AATGTCAAGGCCCTTGATCTGCTCCACGCCGCGCTGTTCACCAACTACGACAGCATGTG 240
OY	241	cagcgcggtgttcccaaggaaaccccccaagaaagtgctccctcggaaatgtgttaccacata 300
Db	241	CAGCGCGGTGTTTCCAAAGAAACCCCAAGAACGTGTCTCTGGAAATGTTGGTCACTTACATA 300
OY	301	gacatcgacgaagtgtgaacgcaagatgtgacccaactgtgtgtctgataatcccgatgaga 360
Db	301	GACATCGACGAAGTTGAACGGCAAGCTGACCAACCCACTGCTGATGAATCTCGATGGAGAA 360
OY	361	gacggagggcgctgtgtgcaaacgfcgaanaatgtgacaatatcagaatcaactttgaag 420
Db	361	GACGGAGGACGGTGTGGCAACCGTCACAAATATGACAAATACCAAGATCACTTTGAAG 420
OY	421	tccaaacgaggtcttgaaaccccccaaatcaacatcttcaaacgagcgaagaagtgtgcctatg 480
Db	421	TCCAAACGAGGTCTGGAACCCCCCAATTCACACTTTCAACGGGAGAGAGTGGCTGATG 480
OY	481	gcccgaaccccaagtgaaaccttaagcccaagaatggcaacttcggttgaatgtccttcgaagcgt 540
Db	481	GCCGGAACCCAGGTGAACCTTCAGCCACGATGCGCCACTTCCGGTGGATGCTCTCACACCGTG 540
OY	541	tacaagcgctacttggaactaacaatgtcaactgtggcccaagaagaagaagaagtcgtcgaag 600
Db	541	TACAGCGCTACTTGGAACACTCAATGCTCAACTGTGGCCCAAGACGAAGAGAGCTGCAAG 600
OY	601	ttgaagaatcgagctcctctggggccttgaagtgcttcttcgcggagaaagcgaagcgaagaa 660
Db	601	TTGAAGATTCGGGTCTCTGGGGCCTGAAGGTGCTGCTGCGGAGAAAGCGACGGCGAGAGA 660
OY	661	gaatccctttgaacacagacgaacctgttcaatgtaaccggaagtgggaaatcgttgatcgcga 720
Db	661	GAGTCCCTTTGACCAAGACGACCTGTTCAAGTCACACCGGAGTGGGAATGATGTGATCTCGCA 720

Qy 721 gccacatttgcagtcagagactaactacggtacatgagtaactctgagcgctcaagcgg 780  
 |||||||  
 Db 721 gccacatttgcagtcagagactaactacggtacatgagtaactctgagcgctcaagcgg 780  
 |||||||  
 Qy 781 cgtctctcatgtatcacgagcggtcatctaacacccgctctcgtcatcgtcatctgccc 840  
 |||||||  
 Db 781 cgtctctcatgtatcacgagcggtcatctaacacccgctctcgtcatcgtcatctgccc 840  
 |||||||  
 Qy 841 ctctcagcctcttgcgctctctccacatgagcgagaaagatcatgatacaagcgctg 900  
 |||||||  
 Db 841 ctctcagcctcttgcgctctctccacatgagcgagaaagatcatgatacaagcgctg 900  
 |||||||  
 Qy 901 ctatcatcgtatgacgcgcctctctcatgtactctgcccagctcctcgcaagtgcttcc 960  
 |||||||  
 Db 901 ctatcatcgtatgacgcgcctctctcatgtactctgcccagctcctcgcaagtgcttcc 960  
 |||||||  
 Qy 961 acaatactcactgtgtatcattcttctacagcaacagcagctgtatctctgagcgttcc 1020  
 |||||||  
 Db 961 acaatactcactgtgtatcattcttctacagcaacagcagctgtatctctgagcgttcc 1020  
 |||||||  
 Qy 1021 acaatctcagagttcttaattcttctgacttgcacagcaagcaagcgagcgctgccc 1080  
 |||||||  
 Db 1021 acaatctcagagttcttaattcttctgacttgcacagcaagcaagcgagcgctgccc 1080  
 |||||||  
 Qy 1081 gaagcgctgagaaagctgctgcacagcgacactgagcgagtgctgctctcgtgcttcc 1140  
 |||||||  
 Db 1081 gaagcgctgagaaagctgctgcacagcgacactgagcgagtgctgctctcgtgcttcc 1140  
 |||||||  
 Qy 1141 agcaacacgtgagagtcgacagcgaggaagaaagatgagcagacccgctacagag 1200  
 |||||||  
 Db 1141 agcaacacgtgagagtcgacagcgaggaagaaagatgagcagacccgctacagag 1200  
 |||||||  
 Qy 1201 gaagcgagagagagagagtcgacagtcgctgagctacacacacagcgagtgccgccc 1260  
 |||||||  
 Db 1201 gaagcgagagagagagagtcgacagtcgctgagctacacacacagcgagtgccgccc 1260  
 |||||||  
 Qy 1261 aagcgcaacagagtcgactgagcgctgctgcacacgcgcgtgagcgcgcatcttctgct 1320  
 |||||||  
 Db 1261 aagcgcaacagagtcgactgagcgctgctgcacacgcgcgtgagcgcgcatcttctgct 1320  
 |||||||  
 Qy 1321 tccctcagcctgagcctctcctcatcttgcagcaatgagtgctcgtgtagagagtgctcaga 1380  
 |||||||  
 Db 1321 tccctcagcctgagcctctcctcatcttgcagcaatgagtgctcgtgtagagagtgctcaga 1380  
 |||||||  
 Qy 1381 ctcaagcgacatcccaagcgagcgagctctgaaactagtttgctgcttgcagtttcat 1440  
 |||||||  
 Db 1381 ctcaagcgacatcccaagcgagcgagctctgaaactagtttgctgcttgcagtttcat 1440  
 |||||||  
 Qy 1441 gtaattatgtgtgtgagactataatttataatgatgagacctcgatgagataaag 1500  
 |||||||  
 Db 1441 gtaattatgtgtgtgagactataatttataatgatgagacctcgatgagataaag 1500  
 |||||||  
 Qy 1501 gaactctgcgaatgtctgtcttacaataaataaataa 1539  
 |||||||  
 Db 1501 gaactctgcgaatgtctgtcttacaataaataaataa 1539  
 |||||||

RESULT 2  
 AY005148 1421 bp mRNA linear INV 01-AUG-2001  
 LOCUS Drosophila melanogaster nicotinic acetylcholine receptor beta  
 DEFINITION subunit mRNA, complete cds.  
 ACCESSION AY005148  
 VERSION AY005148.1 GI:15072330  
 KEYWORDS  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;  
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 1421)  
 AUTHORS Grauso M, and Sattelle D.B.  
 TITLE Drosophila melanogaster nicotinic acetylcholine receptor beta3

JOURNAL subunit (nacrbeta-21C gene)  
 REFERENCE 2 (bases 1 to 1421)  
 AUTHORS Grauso M, and Sattelle D.B.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-JUL-2000) MRC-FGU Human Anatomy and Genetics,  
 University of Oxford, South Parks Road, Oxford OX1 3QX, UK

FEATURES  
 source  
 1..1421  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /chromosome="2"  
 /map="2L; 21C5-21C5"  
 /dev\_stage="embryo"  
 45..1370  
 /note="nacrbeta-21C; ion channel; transmembrane  
 neurotransmitter"  
 /codon\_start=1  
 /product="nicotinic acetylcholine receptor beta subunit"  
 /protein\_id="AF89090.1"  
 /db\_xref="GI:15072331"

CDS  
 463 a 389 g 258 t

BASE COUNT 311 a 463 c 389 g 258 t  
 ORIGIN

Query Match 79.3%; Score 1220; DB 3; Length 1421;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1370; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 14 gtaattcgaagaaagcgagcgctgaaacaaatgagagactcccaagataaagcaccag 73  
 |||||||  
 Db 16 gtaattcgaagaaagcgagcgctgaaacaaatgagagactcccaagataaagcaccag 75  
 |||||||  
 Qy 74 ttctcagtcctgagactgcaactactgtgcaaatgactaagtgagtgctcttctttagggc 133  
 |||||||  
 Db 76 ttctcagtcctgagactgcaactactgtgcaaatgactaagtgagtgctcttctttagggc 135  
 |||||||  
 Qy 134 tgacttcgtgcaagcgagcgcaactgcaacgcgcgagcccaagaacgcaatgtcaaggccc 193  
 |||||||  
 Db 136 tgacttcgtgcaagcgagcgcaactgcaacgcgcgagcccaagaacgcaatgtcaaggccc 195  
 |||||||  
 Qy 194 tggatgccttcgaagcgagcgctgttcaagcaactcgaacgagatgtcaagcggtgttcc 253  
 |||||||  
 Db 196 tggatgccttcgaagcgagcgctgttcaagcaactcgaacgagatgtcaagcggtgttcc 255  
 |||||||  
 Qy 254 aaggaacccccaagaaagcttcccttggaataatggtgtacactacatagacatcgagagt 313  
 |||||||  
 Db 256 aaggaacccccaagaaagcttcccttggaataatggtgtacactacatagacatcgagagt 315  
 |||||||  
 Qy 314 tgaagcgcaagctgagcaaccccaactgctggtgaattccgagtggagagagagcgcg 373  
 |||||||  
 Db 316 tgaagcgcaagctgagcaaccccaactgctggtgaattccgagtggagagagagcgcg 375  
 |||||||  
 Qy 374 ttgtggaacgctgaacatagacaacatcagcagatcaacttgaatgcagagagtgct 433  
 |||||||  
 Db 376 ttgtggaacgctgaacatagacaacatcagcagatcaacttgaatgcagagagtgct 435  
 |||||||  
 Qy 434 ggaacccccaatcacactcttcaacgagcagaaggtggcgtgatgagcgaacccagag 493  
 |||||||  
 Db 436 ggaacccccaatcacactcttcaacgagcagaaggtggcgtgatgagcgaacccagag 495  
 |||||||  
 Qy 494 tgacctcgaagcagatgagcaactcctcggtgagatgctcgaagcggtgtacaagcctact 553  
 |||||||  
 Db 496 tgacctcgaagcagatgagcaactcctcggtgagatgctcgaagcggtgtacaagcctact 555  
 |||||||

QY	554	ggaactcaacatgctcaacttggcccaacgaacagacagctgcgaatttgaagatcggt	613
Db	556	GGGAATCAACATGCTCAACTTGCCCCACGACAGACAGACTGCAAGTTGAAGATCGGCT	615
QY	614	ccttggcgccgaagtgctgcctccgggaacagcagcggcggagagagatcccttgacc	673
Db	616	CCTGGGGCCTGAAGTGCCTCTGCGCGGAAGACGACAGCGGAGAGAGTCCCTTGACC	675
QY	674	acgaagaccctgttgaagcaccgggttggaaatctgttgaccgagagccacttgttca	733
Db	676	ACGAGACACTGGTTAGTACCGGGATGGGAATGTGGACTGGGAGGCCACTTTGTGA	735
QY	734	gtcaaggaactactaagcgtctacaatgagtaactcttgaagcgtctaagcggcctccatagt	793
Db	736	GTCAGAGACTACTACGGGTACATGATGAGTACACTCTGACGGCTCAGCGGGCCTCCATATG	795
QY	794	acaagcgccgtatctcaacacacccggctcctgcatcgtatccttgaccttcaagccctt	853
Db	796	ACACGGCCGTCAATTCACACACCCGGGCTCTGATGTATCTTGAGCCCTCAGGCTTCT	855
QY	854	ggctcccccacacatggcggcggagaagatcatgatcaagcgccctgcatactcgtga	913
Db	856	GGCTGCTCTCCCAATGGGGCGGCGAAGATATGATACGGCCTGCTCATATCGTGA	915
QY	914	tcggcgccctccctaactgtacttcggccagctcctgcagtgctgttccaacaaatactccac	973
Db	916	TGCGCGCCTTCTCAATGACTTCGCGGACGCTCCGACTCCGAGTGTGTCACCAATACTCCAC	975
QY	974	ttgtgtatattcttcaagcacaacagccctgcgttatctgaagcttccacacatcgttcgaag	1033
Db	976	TTGTGTATATCTTCTACAGCACACACCTGCTGTATCTGAGCTCTTCACCATGCTGGAGG	1035
QY	1034	ttctaagctctgtaaccttgcacacagcagaacaaagaagcgccctgcggagagcgctcgaaga	1093
Db	1036	TTCTAGTTCTGTGACTCGGCGCACAGGCAAGACAAAGAGGGGCGCTGCGGAGGGCGGTGAGAA	1095
QY	1094	agctgtctcaaggggaacttggcagctggtgctgtctctcggttgttcagacacacttggcg	1153
Db	1096	AGCTGTCTCACGGGACCTGGGGACAGTGGGTGCTCTCTCGGTGTTTCACACACACCGGGG	1155
QY	1154	agtcgacagcgagagaagacccaagaagatggaagagacacccgtacagagagggcggaagac	1213
Db	1156	AGTGCAGAGCGGAGAGAAGACAAAGAGAGTGTGAGACAGACACCCGTACGAGAGGCGGAGAGAC	1215
Db	1216	AGGATCTCAAGTCCGCTGGGCAATCAACACACACGAGGTGCGGGCGCCCAAGGCAACACAT	1275
QY	1214	aggagtcagtcgcgtcgtgggcatcaacacacacccaggtgtccgggcggcgaagcgaacacagt	1273
Db	1274	tcgaactggcgctgtgtggcaccacgcgcgtggagccgcatcttcgtttccctcaagccttg	1333
Db	1276	TGGACTGGGGCGTGTGGGCCACCGCGTGTGAGACCGCATTTCTTTCGTTCTTCAAGCTTGG	1335
QY	1334	cccttcacatcttggccatcaagtgctcgtgttgaagatgctcgaagactcaag	1386
Db	1336	CCTTCCATCTTGCGCATACAGTGTCTCGTGTAGGAGTCTCGAGACTCAAG	1388
RESULT	3		
DME318761			
LOCUS	DME318761	1334 bp	mRNA
DEFINITION	Drosophila melanogaster	mRNA for nicotinic acetylcholine receptor	linear
ACCESSION	AJ18761		
VERSION	AJ18761.1	GI:15130913	
KEYWORDS	nAChRbeta-21C gene; nicotinic acetylcholine receptor		
SOURCE	fruit fly.		
ORGANISM	Drosophila melanogaster		
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;		
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 1334)		
AUTHORS	Lansdel, S.J. and Millar, N.S.		
JOURNAL	unpublished		

REFERENCE	2 (bases 1 to 1334)
AUTHORS	Miller, N. S.
TITLE	Direct Submission
JOURNAL	Submitted (31-JUL-2001) Miller N.S., Department of Pharmacology, University College London, Gower Street, London, WC1E 6BT, UNITED KINGDOM
FEATURES	Location/Qualifiers
source	1. .1334
gene	/organism="Drosophila melanogaster"
	/db_xref="taxon:7227"
	5. .1330
CDS	/gene="nAcRbeta-21C"
	5. .1330
	/gene="nAcRbeta-21C"
	/function="nicotinic acetylcholine receptor subunit"
	/codon_start=1
	/product="nicotinic acetylcholine receptor beta 3 (Dbeta3) subunit"
	/protein_id="Cac48166.1"
	/db_xref="GI:15130914"
	/translation="MTTTPKIKAPVSGPLPLILMLMGMLMGLTFSVPAPATADPK LNRKALDRLHAGLFETVDVDVOPVPGCTPNSLEAVVYRIDELNGKLTTHCMN LRKDERVWQPSQYDNIOTITLTKSSVYPTQITLFPDGGGLAESTQVLSHDGHR WMPAVYATVCELMMLWMPHDKSCRTKTSWGKLVLPENGRKGSLDHDIQSP EEELVDSRAHFVSQDYGYEMETLFTAORSNMTAVLYTPASCIYLIATLAEWLP GGEIMINGLILIVIAEYFMEFADLPLVLSNNPLVIVFSTSLVSTVIEVL LYATGKHRRRLPEALRKLHGLGTWLLVFSTSGEQAKETEMEHYEADEOE SSPFIHNTPEVPAKANQFMDLATVDRISFVFSFLAIFLAIRCSV"
BASE COUNT	285 a 433 c 368 g 248 t
ORIGIN	
Query Match	76.7%; Score 1181; DB 3; Length 1334;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 1331; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
QY	39 aacaaatgcagcagcagcagcccaagaataaagcagccagcttcggtcgtgactgctgccaactact 98
Db	1 AACAAATGACGAGCATCTCCCAAGATAAAGGCACAACTTCCGGCTCTGACATGCACTACT 60
QY	99 gctgcgaatgctaattggtgagatgctctcttaataagagcttgacttcggtgcagagcgcaatgac 158
Db	61 GCTGCAATGCTAATAGGGGATGCTCTTATATGGGGCTGACTTCCGTGCCAGGCGCCACACTGC 120
QY	159 caccgcgagaccccaagaaacgcgaatgtaagcgccctgatacgctccacgcgcggacgtgt 218
Db	121 CACCCGCGAGCCCAAGAACGCCAATGTCAAGGCCCTGGATGGCTCCACGCCGCGCTGT 180
QY	219 cagcagcagcagcagcagcagatgtagcagcggtgtttccaagaagaccaccaagacgtgctcct 278
Db	181 CACGACTACGACACACGATGTGCACGCCGCTGTCCAAAGAACCCCCACACAGATGTCTCT 240
QY	279 ggaatgtgtgtgcacctacatagaaatctgacagagtttgaaacggcgaagctgacacaccactg 338
Db	241 GGAATGTGTGTGACACTCATATAGACATGACAGAGTTGAACGGCAAGTGACACACCACCTG 300
QY	339 ctgagctgaatctccgatagtgagagagcagagcggtgtgtgcaacgcgtcaacatatgacaa 398
Db	301 CTGGCTGAATCTCCGATGTGAGAGAGAGAGAGCGCGCTGTGGCAACCGTCAACAATATGACAA 360
QY	399 catcacagatcacacttgaagctccagcagaggtctcgagaccccccaatcacactctcaa 458
Db	361 CATACGACGATCACTTTGAAGTCCAGGAGGTCTGGACCCCCCAATACACACTTTTAA 420
QY	459 cggcgaacgaaggtgagcctgtatgaccgaaaccacagtgacccctcagccaagatgagccactt 518
Db	421 CGGCGACGAAGGTGGCCCTGATAGTCGCGAAGAACCCAGGTGACCTTGACCCAGTATGACACT 480
QY	519 ccggtgtgagctgctccagcggtgttaacagcgtcactcgtgcaatcaaatgctcaactggtcc 578
Db	481 CCGGTGTGAGCTCTCCAGCGCGTGTACACGGCGCTTACTGCAACATCAACATGCTCAACTGGCC 540
QY	579 ccaacgaacagcagagctccaagtgtgaagaatcgtgctcctcgaggcctgaaagctgactctcc 638

```

|||||
Db 541 CCACGACACAGAGCTGCAAGTTGAAGATCGGCTCTGGGGCTGAAGGTCTCGCTGCC 600
Qy 639 ggaagaacgagcagcgagagagagatcccttgacacgaagacgtgttcaatcacgga 638
Db 601 GGGAACGCGACGCGGAGAGAGTCCCTTGACACAGACGACCTGGTTAGTCACCGGA 660
Qy 699 gtgggaatcgttgactcgcgcgagccacttgcacgtcgaactactaagcgtacatgga 758
Db 661 GTGGGAATCTGGACTCGGAGCCCACTTTGTGACGACGACTACTACGGCTACATGGA 720
Qy 759 gtacacttgcagcgtcgcgcgtccctccatgtaacagcgccgtatcaacacccgc 818
Db 721 GTACACTGTAGCGCTCAGCGGCGCTCTCCATGATACAGCGCGGTATCAACACCGCG 780
Qy 819 gtccgtcatcgtcatccctggcctccagccttcagctcgtcctcccaatggcgaga 878
Db 781 GTCTGTGATGTCATCTGCGCCCTCTCAGCCTTCTGCTCCCTCCACATGGCGGCGGA 840
Qy 879 gaagatcatgataacgcgcctgtcactcatcgtatcgcgcgccttccatgtaactcgc 938
Db 841 GAAGATCATATATCAACGCGCTGCTCATCATGATGCGCGCTTCTCTCATGTAATTGCG 900
Qy 939 ccaagctctgcagagtcgtgtccacaatactcacttgtgtatcttctacagcacag 998
Db 901 CCAAGCTCTGCGCAAGTGTGTCACAACATACTCCACTGTGTATCTTTCACAGCACAG 960
Qy 999 cctgctgtatcgtagcgttcacacacatgcgcagagtcgtatctgtacgtgcacag 1058
Db 961 CTGCTGTATCTAGAGCTCTCCACATCTGAGAGTCTGATCTGTACTGTGGCCACAGG 1020
Qy 1059 caagcacaagaagcgccctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1118
Db 1021 CAAGCACAAGAAGCGCGCTGCGGAGGCGCTGAGAAAGCTTTTCACGCGGCACTGGGAC 1080
Qy 1119 gtgagtcgtcgtcgtcgtgtgttccagcacacttgcgcagtcgcagcgagagaacaa 1178
Db 1081 GTGGCTGCTCTGCTCGGTGTTCCAGACACACGCGGAGTCCAGGCGGAGAAACAAAGA 1140
Qy 1179 gatggcgaacacccgcgaagagagcgagcgagcgagcagatcagtcgcgtggacaa 1238
Db 1141 GATGGAGACACCCCTGACAGAGGCGGAGGAGCAAGTCCAGTCCGTGGGCAATCAA 1200
Qy 1239 ccaacaccgagtgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1298
Db 1201 CCACACCGAGAGTCCGGGCGCCCAAGGCCAACCAATTGATGGGCGCTGCTGCCACCGC 1260
Qy 1299 cgtggacgcgacttccctcgttccctcgaacgtgcgccttccatcttgacatcagtg 1358
Db 1261 CGTGGACCGCATTTCTTCTTCTTCCCTTCCAGCTGCGCTTCTCATTTCTGCGCATCAGGTG 1320
Qy 1359 ctccgtgtaggat 1372
Db 1321 CTCCTGTAGGAT 1334

```

```

RESULT 4
AC004573 85095 bp DNA linear INV 30-JAN-1999
LOCUS Drosophila melanogaster, chromosome 2L, region 21C5-21D1, P1 clone
DEFINITION DS07610, complete sequence.
ACCESSION AC004573
VERSION AC004573.1 GI:4204255
KEYWORDS HTG.
SOURCE Fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 85095)
Celinker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,

```

```

TITLE
JOURNAL Sequencing of Drosophila chromosome 2L, region 21C5-21D1
REFERENCE Unpublished (1998)
AUTHORS 2 (bases 1 to 85095)
Celinker,S.E., George,R.A., Galle,R., Svirskaas,R.R., Hoskins,R.A.,
Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Chavez,C.,
Chew,M., Doyle,C.M., Farfan,D.E., Flanagan,J., Houston,K.A.,
Hummasi,S.R., Karra,K., Kearney,L., Kim,S.H., Lee,B.,
Lomotan,M.A., Mak,J., Mazda,P., Mok,M.S., Moshrefi,A.R.,
Moshrefi,M., Nixon,K., Pacled,J.M., Park,S., Pfeiffer,B., Punch,E.,
Snit,E., Twomey,B., Wan,K.H., Whitehaw,K.R., Yee,A., Zhang,R.,
Zieran,L.L. and Kimmel,B.E.
Direct Submission
JOURNAL Submitted (17-APR-1998) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Jan 30, 1999 this sequence version replaced gi:4164118.
COMMENT Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
For further information about this sequence, including its location
and relationship to other sequences, please visit our location
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdg@fruitfly.berkeley.edu.
P1 library location: 80-26.
FEATURES
location/Qualifiers
1..85095
/organism="Drosophila melanogaster"
/strain="y2: cn bw sp"
/db_xref="taxon:7227"
/chromosome="2L"
/map="21C5-21D1"
/clone="P1 DS07610 (D275)"
/clone.lib="P1 library, partial Sau3a in pNS582tet1Aad10"
BASE COUNT 22289 a 19671 c 20319 g 22816 t
ORIGIN
Query Match 40.8%; Score 628; DB 3; Length 85095;
Best Local Similarity 100.0%; Pred No. 0;
Matches 628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 352 cgaatggaagacgagagcgctgtgtgcaacgclcaacatataacacacacgagatc 411
Db 5406 CGATGGAGACAGCAGGAGCGGTGTGGCAACGTCACATATATGACAAATCAGCAGATC 5465
Qy 412 acttgaatcgcagcagagtcgtgagccccaataacacactcttcaaggcgagaaagt 471
Db 5466 ACTTTGAAGTCCAGCGAGGCTGTGACCCCAATATCACATCTTCAACGCGCAGAAAGT 5525
Qy 472 ggcctgaatgcggaacccagtgaccctcagacagacagtgccacattccggtgagtgct 531
Db 5526 GGCTGTATGCGCGAAGACCCAGGTGACCTCAAGCAGATGACCACTTCCGTTGATGCTCT 5585
Qy 532 ccaagcgtgtacacgagcctactcgaactcaaatcgaactcgaactcgaactcgaactcga 591
Db 5586 CCAAGCGGTATACAGCGGCTACTGGAATCTCAATGCTCAATGCGCCCAAGCAAGAGAG 5645
Qy 592 agctgcaagtgaagatcgcgtccctgggcgtgaagtgctctcgcgcgagaacgcagc 651
Db 5646 AGCTGCAAGTTGAAGATCGGCTCTCTGAGGCTCTGAGTCTCTGCGCGGAGAACGCGACG 5705
Qy 652 ggcgaagaagaatcccttgacacagacgacctggttaacgaccgcgagtggaatcgtg 711
Db 5706 GCGAGGAGAGATCCCTTGACACGACGACCTGTTTACACCGGAGGTGGAAATGATG 5765
Qy 712 gactcgaagccacttgcagtcaggaactaagcgtcacatgagtaactcgtgagc 771

```



mrna

loc13618 .3858 .11098 .11376 .11383 .11475 .19033 .19179  
27475 .27612 .28770 .28903 .29373 .29558 .29704 .29901,  
30371 .30553 .31293 .31473 .32562 .32796 .34065 .34210,  
35933 .36146 .36220 .37083 .37396 .37843 .37909 .38301,  
38363 .38507 .38569 .39613  
/gene="Plc21c"  
/product="PLI4752"  
/db\_xref="FLYBASE:FBann004574"  
/db\_xref="FLYBASE:FBgn0004611"



```
gene
<3618. >39613
/gene="PLC21C"
/feature="CG4574"
/feature="21B7-21C1"
/db_xref="FLYBASE:FBan0004574"
/db_xref="FLYBASE:FBgn0004611"
3618..11383..11475,19033..19179,27475..27612,28770..28903,
29373..29558,29704..29901,30371..30553,31293..31473,
32562..32796,34065..34210,35933..36146,36220..37083,
37396..37843,37909..38301,38363..38507,38569..38802)
/gene="PLC21C"
/feature="PLC21C gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0004574"
/db_xref="FLYBASE:FBgn0004611"
/protein_id="AAF51507.1"
/db_xref="GI:7296215"
/translation="MMSAGTYITSAVSVEPOLADGEKFIKMDSDSGTGTPTVMAVD
AKGFLYVWDONNELDIDIAIRDTQYAKRPADNLRQIVLTGLPDQTEKVT
VCHGSDVFNMTFVNEFCRDIADLTDLIKIYSLAQLNGSAIMFLOKAKTKLQ
VKSGRIPYKNIKLEPAONKEDKRYEKALDYGLSGYDSISVSKFOEDPYNLK
YLTORSEVERLDSIVGNSKRCMSIAOLVEFLNKORDPRLNEILYPANARAKEL
IOYEPKNNAOKGOLSLDGLFRLMGDNPIMAPSKLDLDDMDPMKHYFINSNH
TYLGHQLTGKSSVELYROCFLAGCQVLEDMNGRTEPEVTHGYTEPEIPAKVL
EALAESAFRTSEYVLLSEFENHCNPROQAKIANYCEIFGDMLLDPDLSHPLEPM
LPPPAMLRKRIIKNNKHHHHHHHHKPPAOVGPANNNKLTANSYAKAAQCV
LSAHEGCVTSTANGDVATGCTGAGTAGHAPLOOIROSSKDSGSDSS
EDELSPNTPLPSGNEPPEKAKETAGETSAIYVYVOPHFSFPAEKKNCY
EMSSPEKATLTKRPIEPVYNNKHQISRYRPACTRFDSSNFMQLFMNQCQIYA
LNFQIDLMQNLGIFETNANSGLILKPEFMRSDRDLPESTVDGIDGTSIT
VLSGFLTKRANTEVEVDVYGLPADTVRKKEFKTVRONGMPLDEEPFEKRYVL
PELASIRIAYEGEGKLIHRYLPLVGLCPGRHVMRSEVGPILASLFLCVVVD
IYDDLSNFEALANIKYQSELEKRIOLSVTDEALSGADELSKSCGKELR
PVESLTSKPRHSISAAAMSVTVDRDGRGSDSISYAPSIQHSIOHSDQST
SIRQVSSQEDVDLVAPELKLIDHKSVEKRLMEKELKRDKEKIKIADOK
SEPLEKRRKRAFTNKLYRKLRSKSLNCSLSPHSEPEVEIPACPLDGDSEBBAAD
GEDLAGSSSLDGRDQESRLRSACRETSQYRSLQRYEATYSAEKVLKTSQYQT
KOLKASLDVTEVQHQLOEARNVEKRLATYARDEDELIRMKREVAASVEGVGER
VRLKQFDRRTDELQKODSVRNALAEHRKARQIADKEAESRCSVSNGLVLFHGP
HHHGTGSSSSALSGNNLTLLNDAGAAAGSHSATSIPAKSHNSIAAAEMKT"
complement(11712..12601)
/mRNA
/gene="CG11912"
/product="CT37080"
/db_xref="FLYBASE:FBan0011912"
/db_xref="FLYBASE:FBgn0031248"
/complement(<11712..>12601)
/gene="CG11912"
/map="21B7-21B8"
/db_xref="FLYBASE:FBan0011912"
/db_xref="FLYBASE:FBgn0031248"
/complement(11773..12588)
/gene="CG11912"
/feature="CG11912 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0011912"
/db_xref="FLYBASE:FBgn0031248"
/protein_id="AAF51509.1"
/db_xref="GI:7296217"
/translation="MKQFAVIFALASVATSVPOGPPEGRITNGYEAKGEARYI
VSIQTSNSHFCAGSLDDEVITYTAHCHCTYNGQVAAHSHSTDEBNQIKETNAQ
YVTHENYGGVGNPNDIGLILKEDEAFDLNAAKDSNPNVSAVLSKTFQGSIDTL
YGGWRNNSGLPLNLOKDAIIVDYNECKAALPSNNSLAETNCTHPTPKADSSCMGD
SGGLVSSSSRAELIGIYSGWGTPLCLSTTTPSVYTSVSSFLPMWIDENKKA"
complement(13648..14580)
/mRNA
/gene="CG11911"
/product="CT37076"
/db_xref="FLYBASE:FBan0011911"
/db_xref="FLYBASE:FBgn0031249"
/complement(<13648..>14580)
/gene="CG11911"
/map="21B7-21B8"
/db_xref="FLYBASE:FBan0011911"
/db_xref="FLYBASE:FBgn0031249"
/gene
```

```
CDS
complement(13743..14576)
/gene="CG11911"
/feature="CG11911 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0011911"
/db_xref="FLYBASE:FBgn0031249"
/protein_id="AAF51508.1"
/db_xref="GI:7296216"
/translation="MKLIIVTVLIALVAAAGAKLSDKIAKLVPSFANGFVINGTEAE
PHSAPIVALANVYLKSHIGCTGLINKMIYTAHCHISVEVPGMSIAGLHTRAEVDE
LTOQROVDGRGHEKRYTGCGPYDIALAHVNSFTFENWVOPATLPSRROVGEETHL
YKQGPKSTYESGAKTLQVTTQIOLALYNECKEELDEPSAIIASNSCSSLQSKACH
GDSGGLVEFTNAPSELIGIYSGWGTPLCLNMPSTYTKSAVIDWITNIDSAITKL
F"
complement(149281..49459,49730..50200,50254..50554,
50786..>51230)
/gene="CG11907"
/product="CT33136"
/db_xref="FLYBASE:FBan0011907"
/db_xref="FLYBASE:FBgn0031250"
/complement(<49281..>51230)
/gene="CG11907"
/map="21C1-21C1"
/db_xref="FLYBASE:FBan0011907"
/db_xref="FLYBASE:FBgn0031250"
/complement(149281..49459,49730..50200,50254..50554,
50786..51049)
/gene="CG11907"
/feature="CG11907 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0011907"
/db_xref="FLYBASE:FBgn0031250"
/protein_id="AAF51506.1"
/db_xref="GI:7296214"
/translation="MGTSDTRIVDALDEDEDELITGESSFVAPFAPPODHADOCIM
EQQDEYVLSNEPESRGTFLTYVELDGLIGMTNPNFTEVEDWKYKERRASINT
DDEELTFLQKSFCDALFATISGTFLLALVIGHNSLRTKMGITMILFLFV
TTEGFIINDTMOQEFLLTLLIIVLNTISAATMGALYGVGLPSEFTAVVSGQA
LGIITLALFIIIVLAFDTPNTTAFIEFFVGVGLILCLCYVILAKRPFYVLEG
DKYKIVRAVPSHNRNGSAGLPLEPILROVMKIKYHAISLALYTTTLYSVPAATVL
MOSEGVSVTAMVMVAFALSNYFTNILLIMPKRVKQREKELASIMAAALSCMAV
GSLISVPOYM"
complement(68682..69037,75067..75450,76257..76302,76864..77132,
77259..77995)
/gene="al"
/feature="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/product="CT12963"
/db_xref="FLYBASE:FBan0003935"
/db_xref="FLYBASE:FBgn0000061"
/feature="al"
/gene="al"
/feature="CG3935"
/map="21C1-21C2"
Query Match 40.8%; Score 628; DB 3; Length 302473;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 352 cgatgaggaagaagagcgctgtgtgcaacgcttcaaatatatacaaacatcagcgatc 411
|||||
Db 236373 CGATGGAAGACAGAGGCGCTGTGCAACCGCTCAAAATATACAAATCAGCAGATC 236432
QY 412 accttgagtcgagcgagctggaacccccaatacacatctctcaacgagcgagcgagtc 471
|||||
Db 236433 ACTTTGAAGTCACAGCGAGCTGTGACCCCAATATACATCTTCAAGCGCAGAGAGT 236492
QY 472 ggcctgagtcgagcgagcgagctggaacccccaatacacatctctcaacgagcgagcgagtc 531
|||||
Db 236493 GGCCTGATGCGCAGAAACCGAGTGACCTTCAGCCAGATGCGCAGCTGATGATGCC 236552
QY 532 ccagcgctgacgagcgagctctggaactcaaatctcaacgagcgagcgagcgagcgagc 591
|||||
```



Db 236553 CCAGCCGCTGTACAGGCGCTACTGCGAATCAATGCTCAATGCGCCCAAGACAGACG 236612  
QY 592 agcgcgaagtgaagatcgcgtccttggtgacctgaagtcgtcctgcgggagaacgcgcacg 651  
Db 236613 AGCTGCAAGTTGAAGATCGGCTCTGGGGCTGAAGGTCGTCTGCGGAGAACGCGACG 236672  
QY 652 gcgcgaagaagatgccttcgtacacgaacgaacgttgcattacacgcgaatgggaatcgtg 711  
Db 236673 GCGAGAGAGAGTCCCTTGACACGACGACCTGGTTCACTACCGGAGATGGGAATCTGTG 236732  
QY 712 gactcgcgaagccacatttgcagtcgaagactactacgcgtacatgagtaacctctgacg 771  
Db 236733 GACTCGGAGAGCCACTTGTGACGACGACTACGCGCTACATGAGATGACACTCTGACG 236792  
QY 772 gctcgaagggcgctcctcattgatacggccgctcattacacccgcgctcctgcatgctc 831  
Db 236793 GCTACGGGCGGCTCTCCATGTACACGCGCTCATCATCACACCGCGCTCTGCATCGTC 236852  
QY 832 atcctgcgcctctcagccttcgtgcctccacatgagcgaggaagatcatgac 891  
Db 236853 ATCTGGCGCTCTCAGCGCTTCTGGCTGCTCCACATGGGCGGAGAACATCATGATC 236912  
QY 892 aagcgctgcctcatcctgtagtcgcgccttcctcctcattcctgcgcgcgcgcgcgcgc 951  
Db 236913 AACGCGCTGCTCATCATCATCATGATGCGCGCTTCTCATGTACTTGCAGCCTCTCGCA 236972  
QY 952 gtgcgtccacacatactcctcctgtgg 979  
Db 236973 GTGCTGTCCAACTACTCTCCACTTGTGG 237000

RESULT 8  
AY061622  
LOCUS AY061622 1646 bp mRNA linear INV 08-NOV-2001  
DEFINITION Drosophila melanogaster SD09326 full length cDNA.  
ACCESSION AY061622.1 GI:16769901  
VERSION AY061622.1  
KEYWORDS FLI.CDNA.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscophora; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1646)  
Stapleton, M., Brokstein, P., Hong, L., Abmayan, A., Carlson, J.,  
Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R.,  
Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J.,  
Nunoo, J., Pacle, J., Paragas, V., Park, S., Phoumenavong, S., Wan, K.,  
Yu, C., Lewis, S. E., Rubin, G. M. and Celisner, S.  
Direct Submission

REFERENCE  
AUTHORS  
TITLE Submitted (30-OCT-2001) Berkeley Drosophila Genome Project,  
JOURNAL Lawrence Berkeley National Laboratory, One Cyclotron Road,  
Berkeley, CA 94720, USA  
COMMENT  
Sequence submitted by:  
Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory  
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to  
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,  
Science 2000). The sequence has been subjected to integrity checks  
for sequence accuracy, presence of a polyA tail and contiguity  
within 100 kb in the genome. Thus we believe the sequence to  
reflect accurately this particular cDNA clone. However, there are  
artifacts associated with the generation of cDNA clones that may  
have not been detected in our initial analyses such as internal  
priming, priming from contaminating genomic DNA, retained introns  
due to reverse transcription of unsplined precursor RNAs, and  
reverse transcriptase errors that result in single base changes.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our web site  
(http://fruitfly.berkeley.edu) or send email to  
cnae@fruitfly.berkeley.edu.

FEATURES

Location/Qualifiers

source 1. .1646  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/map="21C5-21C5"  
1. .1646  
/gene="nAcRbeta-21C"  
/note="alignment with genomic scaffold AE003589"  
/db\_xref="FLYBASE:FBgn0031261"  
37. .1020  
/gene="nAcRbeta-21C"  
/note="longest ORF"  
/codon\_start=1  
/db\_xref="FLYBASE:FBgn0031261"  
/product="SD09326p"  
/protein\_id="AAI29170.1"  
/db\_xref="GI:16769902"  
/translation="MTTPKIKAPVSGPLPLLLMGLMGLTSLVPATATADPK  
NANWGLDRLHAGLFTNIDSDVPFVGFTPTLNGMVTYIDIDELNGLTHWMLN  
LRWDERVWQPSQYDNIQTITLKSSEVMTPTLNGDGLMAETQVTLSDHGR  
MMPAYVTAFCELNMLNMPHDKSCRLKIGSMGLKVLPENGTARDESIDHDLNOSP  
EMETVSRALFVSDYRGVMEYTLTQNRSSMTYVITPASTIVLALSAFWLPHM  
GSEKIMINGLILIVIAFLMTFPAQLPLVLSNNTPLVGECSPLPFHOFSSBN"  
BASE COUNT 390 a 506 c 428 g 322 t  
ORIGIN  
Query Match 32.9%; Score 507; DB 3; Length 1646;  
Best Local Similarity 99.1%; Pred. No. 1e-269;  
Matches 957; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 14 gtacatccgaacaaagcgcgctgaaacatgacgaacactcccaagataaagcacacg 73  
Db 8 GTACATCCGAACAAAGCGCGCTGAACAATGACGACGACTCCCAAGATAAGCACACAG 67  
QY 74 ttccggctcgtgacgctgacactactgctgcaaatgtaatggtggatgcttcttaaggac 133  
Db 68 TTTCCGGTCTCGACTGCACTGCACTACTGCTGCAAAATGATGAGGATGCTTCTTATGGGCG 127  
QY 134 tgacttcggtgcgaagcgccactctgcaacgcggcgcccaagaagccaatgtaagccc 193  
Db 128 TGACTTCCGTGCGCAGGCGCAGCTGCAACCGCGGACCCCAAAAGCCCAATGTCAAGGGCC 187  
QY 194 ttgactgcctcgaacgcggcctgtgtcaacactcgaacgaagatgctgaacgcgtgttc 253  
Db 188 TGGATCGCCTCAAGCGGCGCTGTTCAAGAACTGACGACGAGATGTGACGCGGTGTTCC 247  
QY 254 aaggaaccccccaagaacgctgctccctggaatgctggtcaactacatagacatcgaagat 313  
Db 248 AAGGAACCCCAAGCAAGTGTCTCGGAATGATGATGATGATGATGATGATGATGATGATGAT 307  
QY 314 tgaacggaagctgac 373  
Db 308 TGAACGGCAACCTGACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367  
QY 374 ttgtgcaacgctcaacatatgac 433  
Db 368 TGTGCAACCTCAACATATATACAAATCAACGACGACATCACTTGAATGCCAGCAGTCT 427  
QY 434 ggaaccccccaac 493  
Db 428 GGACCCCCCAAAATCACTCTTCAACGGCGACGAAGCGCGCTGATGGCCGAACCCAGG 487  
QY 494 tgaccctcagcgaagatggaac 553  
Db 488 TGACCCTCAGCAAGATGCGCATTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 547  
QY 554 gcgaac 613  
Db 548 GCGAATCAACATCTCACTGCGCCCAACGACGACGACGACGACGACGACGACGACGACGACG 607  
QY 614 ccttggaacgctgacgctcctgctgcaagacgcaacgcaacgcaacgcaacgcaacgcaacg 673  
Db 608 CCTGGGCGCTGAAGTGTCTGCGGAGAACGCGACGCGACGAGAGAGTCTGCTGACG 667

```

QY 674 acgacgacctgtgtcagtcacccgagtgaggaaatcgttgactcgcgagccacttgcga 733
    |||||||
Db 668 ACACAGACCTGTTTCAGTACACCGAGTGGGAATCGTGACCTGGCGAGCCCACTTTGTCA 727
QY 734 gtcaagaaactactcgcgcgtacatagagtaacacttgacgcgtcagcggcgtctcatgt 793
    |||||||
Db 728 GTTAGACACTACAGGCTACATGAGTACACTGTGACGGCCAGCCGCGCTCTTCATGTG 787
QY 794 acacggcctcatctacacacccgcgtctgcacgtcactcgcgcgtctcgaaccttct 853
    |||||||
Db 788 ACACGGCGGTCTACATACACACCGCTCTGCTCATCTGCGCTTTCAGCCCTTCT 847
QY 854 ggcctgcctcccaacatgaggcgcgagaaatcatgatacaacggcctgcgtcatcgtga 913
    |||||||
Db 848 GCGTGCCTCCCAACATGAGGCGGAGAAAGATCATGATCAACGCGCTGCTCATCATGTGA 907
QY 914 tgcgcgcctctccatcgtactctgcgcgcgcgtctgcgcgcgtctgcgcgcgcgt 973
    |||||||
Db 908 TCGCGCTCTTCATGACTGACTGCTGCGCCAGCTCCTGCCAGTGTCCCAACAAATACCTCAC 967
QY 974 ttgtgg 979
    |||||||
Db 968 TTGTGG 973

```

```

RESULT 9
LOCUS Dm159B9T/c 299 bp DNA linear STS 17-APR-1996
DEFINITION D. melanogaster STS determined from European Mapping Project
ACCESSION Z31885
VERSION Z31885.1 GI:469891
KEYWORDS STS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 299)
AUTHORS European Drosophila Mapping Consortium.
TITLE Direct Submission
JOURNAL Submitted (08-APR-1994) Michael Ashburner, Department of Genetics,
    Downing St., Cambridge CB2 3EH, England
2 (bases 1 to 299)
AUTHORS European Drosophila Mapping Consortium.
TITLE Direct Submission
JOURNAL Submitted (15-APR-1996) Michael Ashburner, Department of Genetics,
    Downing St., Cambridge CB2 3EH, England
REMARK Updated comments
COMMENT STS_name = Dm159B9T
        STS_clone_name = 159B9
        STS_from_promoter = T7
        vector_class = cosmid, Lorist 6
        origin_of_clone = Oregon-R
        in_situ_site_primary = 21C
        STS_dbSTS_AC = 4426
        BLAST_program = BLASTN
        database_searched = EMBL
        date_of_search = 08-01-1996
        date_of_search = 08-01-1996
        BLAST_program = BLASTX
        database_searched = SWISSPROT
        date_of_search = 15-12-1995.
FEATURES
    Source
        1..299
        /organism="Drosophila melanogaster"
        /strain="Oregon-R"
        /db_xref="taxon:7227"
        /clone="159B9"
BASE COUNT 59 a 77 c 89 g 74 t
ORIGIN

```

```

Query Match 9.4%; Score 144; DB 11; Length 299;
Best Local Similarity 100.0%; Pred. No. 4,6e-68;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 gtacatccgaaaacaaagcgcgtgaaacaaatgacgacactcccaagataaaggcaccag 73
    |||||||
Db 161 GTACATCCGAACAAAGAGCGCGCTGAACATGACGACGACTCCCAAGATAAAGGCACAG 102
QY 74 ttccgcgtctgtgacgcgcacactactgtcgaatgttaatgggaatgcttattaggagc 133
    |||||||
Db 101 TTTCCGCTCTGACACTGCCACTACTCTGCAATGCTAATGGGAGTGTCTTATAGGGGC 42
QY 134 tgactcgtgcgcagcgcgcactg 157
    |||||||
Db 41 TGACTTCGCTGCCAGCGCCACTG 18

```

```

RESULT 10
LOCUS AC101533 74135 bp DNA linear HTG 23-NOV-2001
DEFINITION Mus musculus clone RP23-191J14, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC101533
VERSION AC101533.1 GI:17060308
KEYWORDS HTG; HTGS_PHASE0.
SOURCE house mouse.
ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 74135)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-191J14
JOURNAL Unpublished
2 (bases 1 to 74135)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
    Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bonkhalter,B.,
    Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
    Choquel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
    Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
    Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
    Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-pierre,N.,
    Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
    Jones,C., Kamat,A., Karatas,A., Kells,C., Laroque,K.,
    Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
    McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J.,
    Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
    Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
    Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
    Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
    Roman,J., Roselli,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
    Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
    Strauss,N., Subramanian,A., Talamas,J., Testafaye,S., Theodore,J.,
    Topham,K., Travers,M., Travis,N., Triggilio,J., Vassiliev,H.,
    Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.D., Young,G.,
    Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
JOURNAL Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
    Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
    Smit,A.F.A. & Green, P. (1996-1997)
    http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

Center: Whitehead Institute/ MIT Center for Genome Research

Genome Center

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

Project Information

Center project name: I16903

Center clone name: 191\_J\_14

\* NOTE: This record contains 92 individual

\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 714: contig of 714 bp in length  
715 814: gap of 100 bp  
815 1512: contig of 698 bp in length  
1513 1612: gap of 100 bp  
1613 2314: contig of 702 bp in length  
2315 2414: gap of 100 bp  
2415 3125: contig of 711 bp in length  
3126 3225: gap of 100 bp  
3226 3934: contig of 709 bp in length  
3935 4034: gap of 100 bp  
4035 4748: contig of 714 bp in length  
4749 4848: gap of 100 bp  
4849 5553: contig of 705 bp in length  
5554 5653: gap of 100 bp  
5654 6354: contig of 701 bp in length  
6355 6454: gap of 100 bp  
6455 7167: contig of 713 bp in length  
7168 7267: gap of 100 bp  
7268 7979: contig of 712 bp in length  
7980 8079: gap of 100 bp  
8080 8798: contig of 719 bp in length  
8799 8898: gap of 100 bp  
8899 9592: contig of 694 bp in length  
9593 9692: gap of 100 bp  
9693 10393: contig of 701 bp in length  
10394 10493: gap of 100 bp  
10494 11210: contig of 717 bp in length  
11211 11310: gap of 100 bp  
11311 12030: contig of 720 bp in length  
12031 12130: gap of 100 bp  
12131 12843: contig of 713 bp in length  
12844 12943: gap of 100 bp  
12944 13662: contig of 719 bp in length  
13663 13762: gap of 100 bp  
13763 14462: contig of 700 bp in length  
14463 14562: gap of 100 bp  
14563 15265: contig of 703 bp in length  
15266 15365: gap of 100 bp  
15366 16077: contig of 712 bp in length  
16078 16177: gap of 100 bp  
16178 16855: contig of 678 bp in length  
16856 16955: gap of 100 bp  
16956 17667: contig of 712 bp in length  
17668 17767: gap of 100 bp  
17768 18464: contig of 697 bp in length  
18465 18564: gap of 100 bp  
18565 19271: contig of 707 bp in length  
19272 19371: gap of 100 bp  
19372 20089: contig of 718 bp in length  
20090 20189: gap of 100 bp  
20190 20914: contig of 725 bp in length  
20915 21014: gap of 100 bp  
21015 21716: contig of 702 bp in length  
21717 21816: gap of 100 bp  
21817 22532: contig of 716 bp in length  
22533 22632: gap of 100 bp  
22633 23339: contig of 707 bp in length  
23340 23439: gap of 100 bp  
23440 24142: contig of 703 bp in length  
24143 24242: gap of 100 bp  
24243 24951: contig of 709 bp in length  
24952 25051: gap of 100 bp  
25052 25762: contig of 711 bp in length

25763 25862: gap of 100 bp  
25863 26576: contig of 714 bp in length  
26577 26766: gap of 100 bp  
26767 27379: contig of 703 bp in length  
27380 27479: gap of 100 bp  
27480 28196: contig of 717 bp in length  
28197 28296: gap of 100 bp  
28297 29018: contig of 722 bp in length  
29019 29118: gap of 100 bp  
29119 29843: contig of 725 bp in length  
29844 29943: gap of 100 bp  
29944 30664: contig of 721 bp in length  
30665 30764: gap of 100 bp  
30765 31484: contig of 720 bp in length  
31485 31584: gap of 100 bp  
31585 32271: contig of 687 bp in length  
32272 32371: gap of 100 bp  
32372 33070: contig of 699 bp in length  
33071 33170: gap of 100 bp  
33171 33864: contig of 714 bp in length  
33865 33984: gap of 100 bp  
33985 34691: contig of 707 bp in length  
34692 34791: gap of 100 bp  
34792 35505: contig of 714 bp in length  
35506 35605: gap of 100 bp  
35606 36310: contig of 705 bp in length  
36311 36410: gap of 100 bp  
36411 37132: contig of 722 bp in length  
37133 37232: gap of 100 bp  
37233 37917: contig of 685 bp in length  
37918 38017: gap of 100 bp  
38018 38725: contig of 708 bp in length  
38726 38825: gap of 100 bp  
38826 39545: contig of 720 bp in length  
39546 39645: gap of 100 bp  
39646 40358: contig of 713 bp in length  
40359 40458: gap of 100 bp  
40459 41171: contig of 713 bp in length  
41172 41271: gap of 100 bp  
41272 41982: contig of 711 bp in length  
41983 42082: gap of 100 bp  
42083 42767: contig of 685 bp in length  
42768 42867: gap of 100 bp  
42868 43571: contig of 704 bp in length  
43572 43671: gap of 100 bp  
43672 44384: contig of 713 bp in length  
44385 44484: gap of 100 bp  
44485 45190: contig of 706 bp in length  
45191 45290: gap of 100 bp  
45291 45987: contig of 697 bp in length  
45988 46087: gap of 100 bp  
46088 46800: contig of 713 bp in length  
46801 46900: gap of 100 bp  
46901 47620: contig of 720 bp in length  
47621 47720: gap of 100 bp  
47721 48427: contig of 707 bp in length  
48428 48527: gap of 100 bp  
48528 49260: contig of 733 bp in length  
49261 49360: gap of 100 bp  
49361 50055: contig of 695 bp in length  
50056 50155: gap of 100 bp  
50156 50862: contig of 707 bp in length  
50863 50962: gap of 100 bp  
50963 51643: contig of 681 bp in length  
51644 51743: gap of 100 bp  
51744 52444: contig of 701 bp in length  
52445 52544: gap of 100 bp  
52545 53233: contig of 689 bp in length  
53234 53333: gap of 100 bp  
53334 54034: contig of 701 bp in length  
54035 54134: gap of 100 bp  
54135 54841: contig of 707 bp in length  
54842 54941: gap of 100 bp



```

* 52826 54080: contig of 1255 bp in length
* 54081 54180: gap of unknown length
* 54181 55801: contig of 1621 bp in length
* 55802 55901: gap of unknown length
* 55902 56988: contig of 1087 bp in length
* 56989 57088: gap of unknown length
* 57089 58770: contig of 1682 bp in length
* 58771 58870: gap of unknown length
* 58871 60598: contig of 1728 bp in length
* 60599 60698: gap of unknown length
* 60699 61734: contig of 1036 bp in length
* 61735 61834: gap of unknown length
* 61835 62911: contig of 1077 bp in length
* 62912 63011: gap of unknown length
* 63012 64026: contig of 1015 bp in length
* 64027 64126: gap of unknown length
* 64127 65264: contig of 1138 bp in length
* 65265 65364: gap of unknown length
* 65365 66665: contig of 1301 bp in length
* 66666 67844: contig of 1079 bp in length
* 67845 67944: gap of unknown length
* 67945 69145: contig of 1201 bp in length
* 69146 69245: gap of unknown length
* 69246 70292: contig of 1047 bp in length
* 70293 70393: gap of unknown length
* 70393 71451: contig of 1058 bp in length
* 71451 71550: gap of unknown length
* 71551 72603: contig of 1053 bp in length
* 72604 72704: gap of unknown length
* 72704 74424: contig of 1721 bp in length
* 74425 74524: gap of unknown length
* 74525 75961: contig of 1437 bp in length
* 75962 76061: gap of unknown length
* 76062 77575: contig of 1514 bp in length.

```

```

FEATURES
source
1. 77575
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-11A10"

```

```

BASE COUNT 19126 a 16992 c 16937 t 20577 3943 others
ORIGIN

```

```

Query Match 1.6%: Score 24; DB 2; Length 77575;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

> 1516 tctgcttcaaaaaaaaaaaaaa 1539
|||||
Db 52890 TCTGCTTACAAAAA 52913

```

```

RESULT 12
AC090939/c 161361 bp DNA linear PRI 20-MAR-2001
LOCUS Homo sapiens chromosome 3 clone RP11-127114 map 3p, complete
DEFINITION sequence.
ACCESSION AC090939 AC026051
VERSION AC090939.1 GI:13384343
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.
1 (bases 1 to 161361)
Song, L., Bao, J., Fan, O., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J.,
Ding, H., Dong, W., Fan, H., Gong, J., Guan, Q., Gu, X., Guo, D.,
Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C.,
Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B.,
Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Niu, Y., Qi, O., Qi, X., Song, L.,
Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J.,
Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X.,

```

```

Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X.,
Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M.,
Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,
Yu, J. and Yang, H.
Chromosome 3p genomic sequence
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 161361)
AUTHORS
Song, L., Bao, J., Bao, O., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J.,
Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D.,
Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C.,
Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B.,
Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Niu, Y., Qi, O., Qi, X., Song, L.,
Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J.,
Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X.,
Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X.,
Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M.,
Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,
Yu, J. and Yang, H.

```

```

TITLE Direct Submission
JOURNAL Submitted (20-MAR-2001) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
COMMENT On Mar 20, 2001 this sequence version replaced gi:9958159.
-----Genome Center
Center:Beijing Center
Website:http://hgsc.igtp.ac.cn
http://www.genomics.org.cn
Contact:hgsc@igtp.ac.cn
-----Project Information
Center project name:1% project
Center clone name: RP11-127114

```

```

-----Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; EP 5% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 163998 bases at least Q40
Consensus quality: 164438 bases at least Q30
Consensus quality: 164509 bases at least Q20
Insert size: 161364; sum-of-contigs
Quality coverage: 12.88x in Q20 bases;sum-of-contigs
-----

```

```

FEATURES
source
1. 161361
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="RP11-127114"

```

```

BASE COUNT 46789 a 36594 c 35116 g 42862 t
ORIGIN

```

```

Query Match 1.6%: Score 24; DB 9; Length 161361;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

> 1516 tctgcttcaaaaaaaaaaaaaa 1539
|||||
Db 115484 TCTGCTTACAAAAA 115461

```

```

RESULT 13
AC026020/c 178622 bp DNA linear HTG 11-APR-2000
LOCUS Homo sapiens clone RP11-625C9, WORKING DRAFT SEQUENCE, 26 unordered
DEFINITION pieces.
ACCESSION AC026020
VERSION AC026020.2 GI:7534095
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens

```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 178622)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens, clone RP11-625C9  
Unpublished  
2 (bases 1 to 178622)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bede, F.,  
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,  
Campopiano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S.,  
Collamore, A., Cooke, P., Dearliano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,  
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lechoczky, J.,  
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McNeeters, R.,  
Melidiri, J., Menus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talmas, J.,  
Testaye, S., Theodore, J., Tittrell, A., Travers, M., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT

Submitted (19-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 11, 2000 this sequence version replaced gi:7263239.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L8574  
Center clone name: 625\_C\_9

----- Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 158834 bases at least Q40  
Consensus quality: 169105 bases at least Q30  
Consensus quality: 173272 bases at least Q20  
Insert size: 189000; agarose-fp  
Insert size: 176122; sum-of-contigs  
Quality coverage: 3.4 in Q20 bases; agarose-fp  
Quality coverage: 3.6 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently  
\* consists of 26 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1098: contig of 1098 bp in length  
\* 1099 1198: gap of 100 bp  
\* 1199 2437: contig of 1239 bp in length  
\* 2438 2537: gap of 100 bp  
\* 2538 3838: contig of 1301 bp in length  
\* 3839 3938: gap of 100 bp  
\* 3939 5158: contig of 1220 bp in length  
\* 5159 5258: gap of 100 bp  
\* 5259 6847: contig of 1589 bp in length  
\* 6848 6947: gap of 100 bp

6948 8346: contig of 1399 bp in length  
\* 8347 8446: gap of 100 bp  
\* 8447 10954: contig of 2508 bp in length  
\* 10955 11054: gap of 100 bp  
\* 11055 13681: contig of 2627 bp in length  
\* 13682 13781: gap of 100 bp  
\* 13782 17774: contig of 3993 bp in length  
\* 17775 17874: gap of 100 bp  
\* 17875 21926: contig of 4052 bp in length  
\* 21927 22026: gap of 100 bp  
\* 22027 25801: contig of 3775 bp in length  
\* 25802 25901: gap of 100 bp  
\* 25902 30370: contig of 4469 bp in length  
\* 30371 30470: gap of 100 bp  
\* 30471 33259: contig of 2789 bp in length  
\* 33260 33359: gap of 100 bp  
\* 33360 36694: contig of 3335 bp in length  
\* 36695 36794: gap of 100 bp  
\* 36795 40643: contig of 3849 bp in length  
\* 40644 40743: gap of 100 bp  
\* 40744 46066: contig of 5323 bp in length  
\* 46067 46166: gap of 100 bp  
\* 46167 51967: contig of 5801 bp in length  
\* 51968 52067: gap of 100 bp  
\* 52068 59476: contig of 7409 bp in length  
\* 59477 59576: gap of 100 bp  
\* 59577 68880: contig of 9304 bp in length  
\* 68881 68980: gap of 100 bp  
\* 68981 78794: contig of 9814 bp in length  
\* 78795 78894: gap of 100 bp  
\* 78895 89939: contig of 11045 bp in length  
\* 89940 90039: gap of 100 bp  
\* 90040 101801: contig of 11762 bp in length  
\* 101802 101901: gap of 100 bp  
\* 101902 116968: contig of 15067 bp in length  
\* 116969 117068: gap of 100 bp  
\* 117069 132078: contig of 15010 bp in length  
\* 132079 132178: gap of 100 bp  
\* 132179 152928: contig of 20750 bp in length  
\* 152929 153028: gap of 100 bp  
\* 153029 178622: contig of 25594 bp in length.

FEATURES  
source  
1..178622  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-625C9"  
/clone.lib="RP11 Human Male BAC"  
1..1098  
/note="assembly-fragment"  
1199..2437  
/note="assembly-fragment"  
2538..3838  
/note="assembly-fragment"  
3939..5158  
/note="assembly-fragment"  
5259..6847  
/note="assembly-fragment"  
6948..8346  
/note="assembly-fragment"  
8447..10954  
/note="assembly-fragment"  
11055..13681  
/note="assembly-fragment"  
13782..17774  
/note="assembly-fragment"  
17875..21926  
/note="assembly-fragment"  
22027..25801  
/note="assembly-fragment"  
25902..30370  
/note="assembly-fragment"  
30471..33259  
/note="assembly-fragment"

```

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: hummquery@sanger.ac.uk
----- Project Information
Center project name: bm36p22
----- Summary Statistics
Assembly program: XGAP; Version 4.5
Sequencing vector: pLsmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 99% of reads

```

```

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: bm215C14
----- Summary Statistics
Assembly program: XAP4: version 4.5
Sequencing vector: plasmid: L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 236582 bases at least Q40
Consensus quality: 237592 bases at least Q30
Consensus quality: 237788 bases at least Q20
Insert size: 238235; sum-of-contigs
Insert size: 240508; 3.8% error; agarose-fp
Quality coverage: 15.3x in Q20 bases; sum-of-contigs
Quality coverage: 15.29x in Q20 bases; agarose-fp

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2002, 10:19:37 : Search time 1762.48 Seconds  
(without alignments)  
11785.560 Million cell updates/sec

Title: US-09-732-680A-1  
Perfect score: 1539  
Sequence: 1 attcgacagcagggatcac.....cttcaaaaaaaaaaaaaa 1539

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 13736207 segs, 674847542 residues

rd size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_iny:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	288	18.7	448	9	AI542842 SD09326.5
2	162	10.5	890	12	CNS0137Y
3	57	3.7	361	9	AM944588 SD09326.3
4	24	1.6	303	10	BI285394 UI-R-DBO-
5	24	1.6	445	9	AI801505
6	24	1.6	475	10	BI399302 UI-R-DBO-
7	24	1.6	520	10	BI385223
8	24	1.6	548	12	AZ958949 2M0226M20
9	24	1.6	1560	11	AK017944 Mus muscu
10	23	1.5	198	10	BG939598
11	23	1.5	347	9	AI535508
12	23	1.5	351	9	AI574766
13	23	1.5	359	10	BF703043
14	23	1.5	360	10	D68664
15	23	1.5	360	10	D69343
16	23	1.5	372	9	AI853953
17	23	1.5	402	9	AI955429

C 18	23	1.5	423	9	AM520985
C 20	23	1.5	447	12	BH593220
C 19	23	1.5	474	9	AM276309
C 21	23	1.5	486	10	BI896248
C 22	23	1.5	493	10	BF401313
C 23	23	1.5	512	10	BE65378
C 24	23	1.5	525	9	AA899747
C 25	23	1.5	572	10	BM428786
C 26	23	1.5	593	10	BI119595
C 27	23	1.5	603	10	BM428787
C 28	23	1.5	621	10	BG436490
C 29	23	1.5	623	9	AM015683
C 30	23	1.5	666	10	BI668117
C 31	22	1.4	75	9	AA615120
C 32	22	1.4	103	10	BG732100
C 33	22	1.4	126	10	BG375749
C 34	22	1.4	139	10	BF057920
C 35	22	1.4	157	10	BM033721
C 36	22	1.4	160	9	AI349175
C 37	22	1.4	174	9	AI251258
C 38	22	1.4	206	10	BM148619
C 39	22	1.4	224	9	AA964311
C 40	22	1.4	254	9	AM822353
C 41	22	1.4	275	10	BM379065
C 42	22	1.4	275	10	BM379104
C 43	22	1.4	282	10	BF511708
C 44	22	1.4	312	9	AJ285166
C 45	22	1.4	313	9	AI454684
C 46	22	1.4	313	10	BF270202
C 47	22	1.4	315	10	BE802075
C 48	22	1.4	326	10	BF557263
C 49	22	1.4	327	10	BM380422
C 50	22	1.4	353	9	AJ285397
C 51	22	1.4	360	9	AM298450
C 52	22	1.4	364	9	AM045460
C 53	22	1.4	388	12	AZ321692
C 54	22	1.4	394	9	AM298730
C 55	22	1.4	411	9	AM532526
C 56	22	1.4	413	9	BE103145
C 57	22	1.4	416	10	BE992989
C 58	22	1.4	419	9	AM738838
C 59	22	1.4	434	10	BE990134
C 60	22	1.4	436	9	AA052898
C 61	22	1.4	439	9	AM520090
C 62	22	1.4	444	10	BI302642
C 63	22	1.4	466	9	AI840647
C 64	22	1.4	475	9	BE110024
C 65	22	1.4	480	9	AI848123
C 66	22	1.4	511	9	AA964637
C 67	22	1.4	532	12	AO589092
C 68	22	1.4	535	10	BE426242
C 69	22	1.4	548	10	BI133496
C 70	22	1.4	554	12	AZ825313
C 71	22	1.4	557	10	BG841202
C 72	22	1.4	568	9	AUI65548
C 73	22	1.4	581	10	BM382438
C 74	22	1.4	611	12	AZ290954
C 75	22	1.4	632	9	AM327740
C 76	22	1.4	661	10	BM380561
C 77	22	1.4	676	10	BM379176
C 78	22	1.4	680	9	AU067768
C 79	22	1.4	681	10	BF635116
C 80	22	1.4	857	10	BE676634
C 81	21	1.4	73	10	BG621583
C 82	21	1.4	116	10	BG362096
C 83	21	1.4	161	10	BG272467
C 84	21	1.4	169	12	B80272
C 85	21	1.4	176	9	BE043547
C 86	21	1.4	187	9	AI349627
C 87	21	1.4	188	9	AI250383
C 88	21	1.4	190	9	AI349135
C 89	21	1.4	207	9	AI581229
C 90	21	1.4	210	9	BB609102

AM520985	UI-R-BOO-
BH593220	BOHRM35TF
AM276309	xrl10a10.x
BI896248	td54n06.x
BF401313	UI-R-CAO-
BE65378	UI-M-BHO-
AA899747	UI-R-EO-d
BM428786	952023D10
BI119595	BM428787
BM428787	952023D10
BG436490	602509162
AM015683	UI-H-BIOP
BI668117	603295951
AA615120	vn02b09.r
BG732100	ps20c04.y
BG375749	UI-R-CSO-
BF057920	7176h08.y
BM033721	kh69401.y
AI349175	ta73d10.x
AI251258	qv36q03.x
BM148619	TCARP2D63
AA964311	UI-R-CO-g
AM822353	uq18f09.x
BM379065	MEST573-C
BM379104	MEST573-C
BF511708	UI-H-BI4-
AJ285166	4A3B-AAG-
AI454684	UI-R-BT0-
BF270202	GA_EB000
BE802075	sr23q07.y
BF557263	UI-R-CO-g
BM380422	MEST519-F
AJ285397	4A3B-AAT-
AM298450	UI-H-BW0-
AM045460	UI-M-BH1-
AZ321692	1M0042C24
AM298730	UI-H-BW0-
AM532526	UI-R-BSO-
BE103145	UI-R-BT1-
BE992989	UI-M-BZ1-
AM738838	gpd3b01.y
BE990134	UI-M-BZ1-
AA052898	z171903.s
AM520090	660056H04
BI302642	UI-R-DMO-
AI840647	UI-M-AM0-
BE110024	UI-R-CAO-
AI848123	UI-M-AP1-
AA964637	UI-R-CO-h
AO589092	CITR-EI-
BE426242	WHE0329_A
BI133496	UI-M-BH3-
AZ825313	2M0100L24
BG841202	MEST18-C0
AUI65548	AUI65548
BM382438	MEST549-F
AZ290954	RPCI-23-1
AM327740	d101f10.x
BM380561	MEST521-F
BM379176	MEST500-G
AU067768	AU067768
BF635116	NU079H02D
BE676634	602086484
BG621583	602617180
BG362096	gbd5b06.y
BG272467	nah3b12.y
B80272	CIT-HSP-204
BE043547	ht44007.y
AI349627	ta74h09.x
AI250383	qx23f03.x
AI349135	ta51h09.x
AI581229	td70h08.x
BB609102	BB609102

C 91	21	1.4	227	10	BM390422	BM390422	UI-R-CNI-	164	20	1.3	102	10	BI682371	BI682371	463436	MA
C 92	21	1.4	231	10	BI341453	BI341453	368885	MA	C 165	20	1.3	116	10	BE634750	BE634750	uvb6e12.x
C 93	21	1.4	240	9	AM314351	AM314351	10381	MAR	C 166	20	1.3	118	9	AI463121	AI463121	uc44c03.x
C 94	21	1.4	254	9	AV735874	AV735874	AV735874		C 167	20	1.3	119	10	BI651387	BI651387	na137b07.
C 95	21	1.4	257	9	AM071389	AM071389	x664g01.x		C 168	20	1.3	119	10	BE511759	BE511759	UI-R-BT4-
C 96	21	1.4	301	9	AI089807	AI089807	qal1b08.x		C 169	20	1.3	121	10	BI293609	BI293609	UI-R-DK0-
C 97	21	1.4	301	9	AM294743	AM294743	UI-H-BW0-		C 170	20	1.3	125	10	BI618756	BI618756	RS76698 A
C 98	21	1.4	303	9	AI562632	AI562632	TENS2618		C 171	20	1.3	126	10	BE955288	BE955288	UI-M-BH4-
C 99	21	1.4	305	9	AI159836	AI159836	qB51a08.x		C 172	20	1.3	127	10	BE583559	BE583559	3-7D-HA P
C 100	21	1.4	308	10	BE377237	BE377237	UI-R-CU0-		C 173	20	1.3	133	9	AM021567	AM021567	dE25d03.Y
C 101	21	1.4	314	9	BE110420	BE110420	UI-R-BJ1-		C 174	20	1.3	137	9	BE116928	BE116928	UI-R-B51-
C 102	21	1.4	320	10	Z36865	Z36865	HHEA693	Att	C 175	20	1.3	138	9	AM139339	AM139339	UI-H-B11-
C 103	21	1.4	325	12	A2693539	A2693539	ZM0232C16		C 176	20	1.3	156	10	BE990522	BE990522	UI-M-B21-
C 104	21	1.4	327	10	BE509458	BE509458	UI-H-B14-		C 177	20	1.3	160	10	BG940656	BG940656	ax07a01.Y
C 105	21	1.4	331	9	AI535265	AI535265	UI-R-C3-S		C 178	20	1.3	162	9	AI792550	AI792550	q171g02.Y
C 106	21	1.4	362	9	AA836519	AA836519	od22d03.S		C 179	20	1.3	163	9	AM397667	AM397667	sg82d09.Y
C 107	21	1.4	368	10	BF459090	BF459090	UI-M-B21-		C 180	20	1.3	164	9	AI136841	AI136841	UI-R-C2D-
C 108	21	1.4	369	9	AI145074	AI145074	UI-R-BT0-		C 181	20	1.3	169	9	AI305981	AI305981	qW48g11.x
C 109	21	1.4	371	9	AJ285066	AJ285066	AA3B-AAF-		C 182	20	1.3	171	9	BI474953	BI474953	IP35c01.x
C 110	21	1.4	381	9	AI954384	AI954384	wX95h06.x		C 183	20	1.3	171	9	AA925693	AA925693	UI-R-A1-e
C 111	21	1.4	393	10	BE557416	BE557416	FI03b12.Y		C 184	20	1.3	171	9	AI436532	AI436532	th52c01.x
C 112	21	1.4	394	10	BF408009	BF408009	UI-R-BJ2-		C 185	20	1.3	171	10	BM385896	BM385896	UI-R-DM1-
C 113	21	1.4	404	10	BE333206	BE333206	uT01g01.Y		C 186	20	1.3	174	9	AI311029	AI311029	ta45e04.x
C 114	21	1.4	417	9	AM047241	AM047241	UI-M-BH1-		C 187	20	1.3	175	9	AI399123	AI399123	NCM13A12T
C 115	21	1.4	424	9	AI699030	AI699030	TX74D08.x		C 188	20	1.3	175	9	AI613361	AI613361	LT81c06.x
C 116	21	1.4	426	10	BI503994	BI503994	BI170024B		C 189	20	1.3	176	9	AI251527	AI251527	qV28a08.Y
C 117	21	1.4	439	9	AI070882	AI070882	UI-R-C2-m		C 190	20	1.3	178	9	AI254848	AI254848	qV36a07.x
C 118	21	1.4	441	9	AI201294	AI201294	qF65h12.x		C 191	20	1.3	178	9	AI311432	AI311432	ta45c01.x
C 119	21	1.4	454	10	BM388354	BM388354	UI-R-CNI-		C 192	20	1.3	179	9	AI254370	AI254370	qV33b03.x
C 120	21	1.4	456	9	AA754029	AA754029	97BS0497		C 193	20	1.3	179	10	BE981135	BE981135	UI-M-CG0-
C 121	21	1.4	464	9	BE026336	BE026336	dB37d04.Y		C 194	20	1.3	182	9	BE042120	BE042120	hJ15a04.Y
C 122	21	1.4	474	9	AI365649	AI365649	MCBA01F07		C 195	20	1.3	182	10	BF986605	BF986605	CM0-GN10
C 123	21	1.4	474	9	AM067248	AM067248	683019F10		C 196	20	1.3	182	10	BF994984	BF994984	CM0-GN10
C 124	21	1.4	480	10	BI295392	BI295392	UI-R-DK0-		C 197	20	1.3	183	9	AI366174	AI366174	CM0-GN10
C 125	21	1.4	487	9	AI097287	AI097287	qB88f12.x		C 198	20	1.3	184	9	AI313345	AI313345	ta77h04.x
C 126	21	1.4	492	10	BE409253	BE409253	UI-R-BT1-		C 199	20	1.3	184	9	AV739755	AV739755	AV739755
C 127	21	1.4	498	10	BM385795	BM385795	UI-R-CNI-		C 200	20	1.3	184	10	BF250279	BF250279	pa97b06.Y
C 128	21	1.4	501	9	AM061064	AM061064	UI-M-BH1-		C 201	20	1.3	185	9	AI308397	AI308397	LT43d09.x
C 129	21	1.4	503	9	AM951298	AM951298	EST363368		C 202	20	1.3	185	9	AA145807	AA145807	mt32d10.X
C 130	21	1.4	524	9	AL384959	AL384959	MBC25E07		C 203	20	1.3	185	9	AV735467	AV735467	AV735467
C 131	21	1.4	526	10	BI796115	BI796115	H038F05 E		C 204	20	1.3	186	9	AM204754	AM204754	UI-H-B11-
C 132	21	1.4	528	9	AL368733	AL368733	MBA26D02		C 205	20	1.3	187	9	AM297442	AM297442	UI-H-BW0-
C 133	21	1.4	568	10	BI274264	BI274264	UI-R-CW0-		C 206	20	1.3	188	9	AM654270	AM654270	103603 MA
C 134	21	1.4	585	10	BE268694	BE268694	1000208D0		C 207	20	1.3	189	9	AM302047	AM302047	xr52f08.x
C 135	21	1.4	586	10	BM379871	BM379871	ME5711-G		C 208	20	1.3	190	9	BE114657	BE114657	UI-R-BJ1-
C 136	21	1.4	600	9	AA898269	AA898269	NCC2A12T7		C 209	20	1.3	191	9	AI029192	AI029192	UI-R-C0-1
C 137	21	1.4	602	9	AI722145	AI722145	f019g08.Y		C 210	20	1.3	193	9	AI312498	AI312498	ta91g10.x
C 138	21	1.4	608	10	BI290555	BI290555	UI-R-DK0-		C 211	20	1.3	193	10	BI945735	BI945735	sc66e07.Y
C 139	21	1.4	626	9	AV694025	AV694025	AV654025		C 212	20	1.3	195	9	AM120595	AM120595	UI-M-BH2
C 140	21	1.4	634	10	BE575184	BE575184	602597719		C 213	20	1.3	198	10	BI321192	BI321192	sa148f05.
C 141	21	1.4	637	10	BE503105	BE503105	602551089		C 214	20	1.3	201	9	AI836404	AI836404	UI-M-AP0-
C 142	21	1.4	640	10	BI798154	BI798154	H099B04 E		C 215	20	1.3	203	10	BE706513	BE706513	280873 MA
C 143	21	1.4	644	9	AV685350	AV685350	AV685350		C 216	20	1.3	204	10	BI280016	BI280016	UI-R-DD0-
C 144	21	1.4	669	10	BF684272	BF684272	602141190		C 217	20	1.3	205	9	AI192763	AI192763	q62g12.x
C 145	21	1.4	689	9	AU092678	AU092678	AU092678		C 218	20	1.3	206	9	AI611528	AI611528	LT65f01.x
C 146	21	1.4	699	10	BE731852	BE731852	601567032		C 219	20	1.3	207	9	AA490702	AA490702	UI-M-BH3-
C 147	21	1.4	700	9	AL501893	AL501893	AL501893		C 220	20	1.3	207	10	BF019259	BF019259	fy03a11.Y
C 148	21	1.4	702	10	BE254472	BE254472	601111562		C 221	20	1.3	211	10	BM026482	BM026482	uF05a09.x
C 149	21	1.4	711	10	BE6840956	BE6840956	MEST13-H0		C 222	20	1.3	215	9	AI366090	AI366090	h017e08.x
C 150	21	1.4	715	10	BE267019	BE267019	601192680		C 223	20	1.3	215	9	AM779032	AM779032	h017e08.x
C 151	21	1.4	723	10	BE260856	BE260856	601151266		C 224	20	1.3	218	10	BE162062	BE162062	dc45c02.Y
C 152	21	1.4	757	10	BI853586	BI853586	603380709		C 225	20	1.3	219	9	AI932631	AI932631	wo05e05.x
C 153	21	1.4	760	10	BE514673	BE514673	60131737		C 226	20	1.3	219	9	AI986680	AI986680	rs76698 A
C 154	21	1.4	777	12	BH510568	BH510568	BOHR05TE		C 227	20	1.3	221	9	BE049186	BE049186	xr73e08.x
C 155	21	1.4	797	10	BE334782	BE334782	602460979		C 228	20	1.3	221	9	AA494912	AA494912	fa12g04.s
C 156	21	1.4	922	12	CMS05T00	CMS05T00	AL339129	Tetraodon	C 229	20	1.3	225	9	AI252894	AI252894	qv31c12.x
C 157	21	1.4	937	10	BM015808	BM015808	603642074		C 230	20	1.3	225	10	BE953771	BE953771	UI-M-CC1-
C 158	21	1.4	1272	12	AG082552	AG082552	Pan t10g1		C 231	20	1.3	226	9	AI251068	AI251068	qx76a08.x
C 159	21	1.4	2747	11	BC013664	BC013664	Mus muscu		C 232	20	1.3	226	9	AM078837	AM078837	xb17h09.x
C 160	20	1.3	48	10	BM144540	BM144540	TCAP1D11		C 233	20	1.3	227	10	BE139372	BE139372	xr69f07.x
C 161	20	1.3	64	9	AI254470	AI254470	qV47h04.x		C 234	20	1.3	227	10	BF932565	BF932565	UI-R-CA0-
C 162	20	1.3	98	9	AM311302	AM311302	sg35b11.Y		C 235	20	1.3	228	9	AI986114	AI986114	wR06e03.x
C 163	20	1.3	102	9	AM457474	AM457474	UI-M-BH3-		C 236	20	1.3	232	10	BE684622	BE684622	186255 MA

C 237	20	1.3	234	9	AA595937	nn6bd05.s	310	20	1.3	300	10	BE584667
C 238	20	1.3	235	9	AU053650		311	20	1.3	302	9	AV737425
C 239	20	1.3	236	9	AA433561	UT-R-BJ0P	C 312	20	1.3	303	9	AI452006
C 240	20	1.3	237	9	AM304297	xy58d07.x	C 313	20	1.3	303	9	AM904833
C 241	20	1.3	240	9	AI361630		C 314	20	1.3	304	9	AA774031
C 242	20	1.3	241	10	BG362300	qy86c04.x	C 315	20	1.3	305	9	AI850824
C 243	20	1.3	242	10	BM087542	BM087542 500327 MA	C 316	20	1.3	305	10	BG880093
C 244	20	1.3	243	9	AI839779	AI839779 UT-M-AM0-	C 317	20	1.3	306	9	AA586909
C 245	20	1.3	245	9	AA260643	AA260643 va12f11.t	C 318	20	1.3	306	10	BF406038
C 246	20	1.3	245	12	AQ040790	AQ040790 CIT-HSP-2	C 319	20	1.3	308	9	AA925564
C 247	20	1.3	247	9	AV741838	AV741838 AV741838	C 320	20	1.3	310	9	AI073243
C 248	20	1.3	247	10	BF469006	BF469006 UT-M-BH3-	C 321	20	1.3	310	9	BB099461
C 249	20	1.3	248	9	AU101315	AU101315 AU101315	C 322	20	1.3	311	9	AI059421
C 250	20	1.3	248	10	BI535631	BI535631 399268 MA	C 323	20	1.3	311	9	BE115146
C 251	20	1.3	248	10	C22109	C22109 M4YA	C 324	20	1.3	312	10	BI535572
C 252	20	1.3	249	10	BE530645	BE530645 M781235TM	C 325	20	1.3	313	10	BE989971
C 253	20	1.3	250	9	AI853431	AI853431 UT-M-BH0-	C 326	20	1.3	314	9	AA555317
C 254	20	1.3	250	10	BE949196	BE949196 UT-M-BH3-	C 327	20	1.3	315	9	AA406310
C 255	20	1.3	252	9	AV373870	AV373870 AV373870	C 328	20	1.3	315	10	BE985968
C 256	20	1.3	253	9	AI763501	AI763501 UT-R-Y0-a	C 329	20	1.3	315	10	BE986834
C 257	20	1.3	253	9	AI856083	AI856083 sg31d07.x	C 330	20	1.3	315	10	BE989720
C 258	20	1.3	253	9	AM284507	AM284507 UT-H-BI2-	C 331	20	1.3	317	10	BM153387
C 259	20	1.3	253	10	BF392335	BF392335 UT-R-CA0-	C 332	20	1.3	318	10	BG385066
C 260	20	1.3	254	10	BG018500	BG018500 daa46d06.	C 333	20	1.3	321	9	AI098578
C 261	20	1.3	255	9	AI310919	AI310919 ta76a05.x	C 334	20	1.3	321	9	AA252799
C 262	20	1.3	255	9	AI312243	AI312243 ta79e10.x	C 335	20	1.3	323	9	AA116810
C 263	20	1.3	257	9	AA478793	AA478793 ryz20d08.s	C 336	20	1.3	323	9	AI714209
C 264	20	1.3	260	10	BM193791	BM193791 TCAPAP1D3	C 337	20	1.3	325	9	BE107137
C 265	20	1.3	260	10	BF049568	BF049568 32DS Amerc	C 338	20	1.3	326	9	AA528617
C 266	20	1.3	262	9	AI249642	AI249642 qx76g10.x	C 339	20	1.3	327	9	AI117576
C 267	20	1.3	263	10	BF388379	BF388379 UT-R-CA1-	C 340	20	1.3	327	9	AW703191
C 268	20	1.3	265	9	AI311422	AI311422 ta45a11.x	C 341	20	1.3	327	10	BI276342
C 269	20	1.3	266	9	AV742051	AV742051 AV742051	C 342	20	1.3	327	10	BE987503
C 270	20	1.3	266	10	R64886	R64886 13390 lambda	C 343	20	1.3	329	9	AW345723
C 271	20	1.3	267	9	AV739381	AV739381 AV739381	C 344	20	1.3	329	10	D22540
C 272	20	1.3	267	10	BF049570	BF049570 98/700 Am	C 345	20	1.3	330	9	AA829834
C 273	20	1.3	268	9	AI583821	AI583821 te73b09.x	C 346	20	1.3	330	9	AI547986
C 274	20	1.3	269	9	AI044518	AI044518 UT-R-C1-K	C 347	20	1.3	330	9	AA489749
C 275	20	1.3	269	10	BM465647	BM465647 AGENCOUNT	C 348	20	1.3	330	10	R39624
C 276	20	1.3	270	9	AM323007	AM323007 u056b11.x	C 349	20	1.3	331	9	AW449165
C 277	20	1.3	270	10	BF415779	BF415779 UT-R-CA0-	C 350	20	1.3	331	10	BI694013
C 278	20	1.3	272	9	AI340518	AI340518 tb30c01.x	C 351	20	1.3	332	9	AI575131
C 279	20	1.3	272	9	AV740587	AV740587 AV740587	C 352	20	1.3	332	9	AA297065
C 280	20	1.3	272	9	BB020648	BB020648 BB020648	C 353	20	1.3	335	10	BG944551
C 281	20	1.3	272	9	BB058442	BB058442 BB058442	C 354	20	1.3	336	9	AI851898
C 282	20	1.3	272	10	BI403748	BI403748 MT-P-CPI-	C 355	20	1.3	336	9	AU066910
C 283	20	1.3	274	9	AV116578	AV116578 AV116578	C 356	20	1.3	336	9	AV742675
C 284	20	1.3	274	10	BI804100	BI804100 H127C01 E	C 357	20	1.3	337	9	AI028959
C 285	20	1.3	275	9	AI711026	AI711026 UT-R-AE1-	C 358	20	1.3	338	9	AI703814
C 286	20	1.3	275	10	BF793539	BF793539 602255056	C 359	20	1.3	338	9	AA434201
C 287	20	1.3	277	9	AL659459	AL659459 AL659459	C 360	20	1.3	339	12	AZ552303
C 288	20	1.3	279	10	BI287759	BI287759 UT-R-CS0S	C 361	20	1.3	340	9	AI844015
C 289	20	1.3	280	9	AI030319	AI030319 UT-R-C0-1	C 362	20	1.3	341	9	AI842121
C 290	20	1.3	280	9	AI311624	AI311624 ta43g06.x	C 363	20	1.3	341	9	AA099119
C 291	20	1.3	282	10	BI007372	BI007372 MRI-RT007	C 364	20	1.3	341	10	BG731992
C 292	20	1.3	284	10	BI688659	BI688659 603311280	C 365	20	1.3	341	10	HI0028
C 293	20	1.3	285	9	AI463565	AI463565 vw63c09.x	C 366	20	1.3	343	9	AI555370
C 294	20	1.3	285	10	BI430970	BI430970 949063E08	C 367	20	1.3	343	9	AM158360
C 295	20	1.3	285	10	C22976	C22976 C22976 Dict	C 368	20	1.3	344	9	AA028165
C 296	20	1.3	288	10	BM279585	BM279585 71/700 Am	C 369	20	1.3	344	9	AI056027
C 297	20	1.3	289	9	AI547649	AI547649 UT-R-C3-S	C 370	20	1.3	344	9	AI705027
C 298	20	1.3	289	9	AI841050	AI841050 UT-M-AM0-	C 371	20	1.3	344	9	AA910018
C 299	20	1.3	289	9	AA565802	AA565802 nk26c07.s	C 372	20	1.3	344	10	BM383300
C 300	20	1.3	291	9	AI611260	AI611260 tu14h11.x	C 373	20	1.3	345	9	AJ280951
C 301	20	1.3	291	9	BE023949	BE023949 sm94f01.y	C 374	20	1.3	346	10	BG361998
C 302	20	1.3	292	9	AA955117	AA955117 UT-R-A1-e	C 375	20	1.3	347	9	AI637031
C 303	20	1.3	292	9	BB070625	BB070625 BB070625	C 376	20	1.3	348	9	AA960497
C 304	20	1.3	293	9	BE115220	BE115220 UT-R-BJ1-	C 377	20	1.3	348	9	AI686698
C 305	20	1.3	295	9	AV743970	AV743970 AV743970	C 378	20	1.3	349	9	AI731642
C 306	20	1.3	295	10	BF467202	BF467202 UT-M-CG0P	C 379	20	1.3	349	10	BF704512
C 307	20	1.3	297	10	BG372285	BG372285 UT-R-CV0-	C 380	20	1.3	350	10	BG734845
C 308	20	1.3	298	9	AA476285	AA476285 w295b09.s	C 381	20	1.3	351	9	AI840625
C 309	20	1.3	298	10	BF074600	BF074600 222056 MA	C 382	20	1.3	351	9	AM597704
C 237	20	1.3	234	9	AA595937	nn6bd05.s	310	20	1.3	300	10	BE584667
C 238	20	1.3	235	9	AU053650		311	20	1.3	302	9	AV737425
C 239	20	1.3	236	9	AA433561	UT-R-BJ0P	C 312	20	1.3	303	9	AI452006
C 240	20	1.3	237	9	AM304297	xy58d07.x	C 313	20	1.3	303	9	AM904833
C 241	20	1.3	240	9	AI361630		C 314	20	1.3	304	9	AA774031
C 242	20	1.3	241	10	BG362300	qy86c04.x	C 315	20	1.3	305	9	AI850824
C 243	20	1.3	242	10	BM087542	BM087542 500327 MA	C 316	20	1.3	305	10	BG880093
C 244	20	1.3	243	9	AI839779	AI839779 UT-M-AM0-	C 317	20	1.3	306	9	AA586909
C 245	20	1.3	245	9	AA260643	AA260643 va12f11.t	C 318	20	1.3	306	10	BF406038
C 246	20	1.3	245	12	AQ040790	AQ040790 CIT-HSP-2	C 319	20	1.3	308	9	AA925564
C 247	20	1.3	247	9	AV741838	AV741838 AV741838	C 320	20	1.3	310	9	AI073243
C 248	20	1.3	247	10	BF469006	BF469006 UT-M-BH3-	C 321	20	1.3	310	9	BB099461
C 249	20	1.3	248	9	AU101315	AU101315 AU101315	C 322	20	1.3	311	9	AI059421
C 250	20	1.3	248	10	BI535631	BI535631 399268 MA	C 323	20	1.3	311	9	BE115146
C 251	20	1.3	248	10	C22109	C22109 M4YA	C 324	20	1.3	312	10	BI535572
C 252	20	1.3	249	10	BE530645	BE530645 M781235TM	C 325	20	1.3	313	10	BE989971
C 253	20	1.3	250	9	AI853431	AI853431 UT-M-BH0-	C 326	20	1.3	314	9	AA555317
C 254	20	1.3	250	10	BE949196	BE949196 UT-M-BH3-	C 327	20	1.3	315	9	AA406310
C 255	20	1.3	252	9	AV373870	AV373870 AV373870	C 328	20	1.3	315	10	BE985968
C 256	20	1.3	253	9	AI763501	AI763501 UT-R-Y0-a	C 329	20	1.3	315	10	BE986834
C 257	20	1.3	253	9	AI856083	AI856083 sg31d07.x	C 330	20	1.3	315	10	BE989720
C 258	20	1.3	253	9	AM284507	AM284507 UT-H-BI2-	C 331	20	1.3	317	10	BM153387
C 259	20	1.3	253	10	BF392335	BF392335 UT-R-CA0-	C 332	20	1.3	318	10	BG385066
C 260	20	1.3	254	10	BG018500	BG018500 daa46d06.	C 333	20	1.3	321	9	AI098578
C 261	20	1.3	255	9	AI310919	AI310919 ta76a05.x	C 334	20	1.3	321	9	AA252799
C 262	20	1.3	255	9	AI312243	AI312243 ta79e10.x	C 335	20	1.3	323	9	AA116810
C 263	20	1.3	257	9	AA478793	AA478793 ryz20d08.s	C 336	20	1.3	323	9	AI714209
C 264	20	1.3	260	10	BM1							

C 383	20	1.3	355	10	BI298246	C 456	20	1.3	393	10	BI703530	BI703530 f691b11.x
C 384	20	1.3	355	10	BE397455	C 457	20	1.3	394	9	A1752895	A1752895 cr02a01.x
C 385	20	1.3	355	10	BE513273	C 458	20	1.3	395	9	A1849514	A1849514 UI-M-BH1-
C 386	20	1.3	357	9	AM121811	C 459	20	1.3	396	9	A1601923	A1601923 UI-R-AD0-
C 387	20	1.3	358	9	A1454747	C 460	20	1.3	396	9	A1955893	A1955893 wt31b04.x
C 388	20	1.3	358	9	A1713509	C 461	20	1.3	396	10	BI808372	BI808372 C008d11.O
C 389	20	1.3	358	9	A1716632	C 462	20	1.3	397	9	A1044054	A1044054 UI-R-Cl-1
C 390	20	1.3	358	9	AV738725	C 463	20	1.3	397	9	A1863117	A1863117 t243e09.x
C 391	20	1.3	358	10	BI942310	C 464	20	1.3	397	9	AM488796	AM488796 UI-M-BH3-
C 392	20	1.3	358	10	BE388140	C 465	20	1.3	397	10	BI538612	BI538612 434014.MA
C 393	20	1.3	359	9	A1556190	C 466	20	1.3	397	12	A2553606	A2553606 RPCI-23-2
C 394	20	1.3	360	10	BE553087	C 467	20	1.3	397	12	A2553606	A2553606 RPCI-23-2
C 395	20	1.3	361	9	AV741090	C 468	20	1.3	400	9	AV741927	AV741927 AV741927
C 396	20	1.3	361	9	AV741092	C 469	20	1.3	400	10	BE042214	BE042214 S093b08.y
C 397	20	1.3	361	10	BI118928	C 470	20	1.3	400	10	BE225081	BE225081 uz42g02.y
C 398	20	1.3	362	9	AV739589	C 471	20	1.3	402	9	AM015245	AM015245 UI-H-B10P
C 399	20	1.3	362	9	AM015593	C 472	20	1.3	402	10	BI142244	BI142244 SDI6292.5
C 400	20	1.3	362	10	BE671410	C 473	20	1.3	403	9	AL502654	AL502654 AL502654
C 401	20	1.3	362	10	BI404266	C 474	20	1.3	404	9	AA614596	AA614596 np50c01.s
C 402	20	1.3	363	9	AM046223	C 475	20	1.3	405	10	BE947234	BE947234 UI-M-BH3-
C 403	20	1.3	363	9	AM060472	C 476	20	1.3	406	9	BE034134	BE034134 MG05D07.M
C 404	20	1.3	364	9	AM523003	C 477	20	1.3	406	9	AA579238	AA579238 n136c08.s
C 405	20	1.3	364	10	BE418467	C 478	20	1.3	406	10	BE788803	BE788803 602110511
C 406	20	1.3	365	10	BI281516	C 479	20	1.3	407	9	AA830974	AA830974 oc62b06.s
C 407	20	1.3	366	9	AM028405	C 480	20	1.3	407	9	AI502510	AI502510 UI-R-C2-n
C 408	20	1.3	366	10	BE990575	C 481	20	1.3	407	9	A1842642	A1842642 UI-M-AQ1-
C 409	20	1.3	366	10	BE512429	C 482	20	1.3	407	9	AM489287	AM489287 UI-M-BH3-
C 410	20	1.3	368	10	BI286153	C 483	20	1.3	408	12	AQ473338	AQ473338 C17BI-E1-
C 411	20	1.3	369	9	AA899709	C 484	20	1.3	409	9	AU031938	AU031938 AU031938
C 412	20	1.3	369	9	A1044563	C 485	20	1.3	409	9	AV737699	AV737699 AV737699
C 413	20	1.3	369	9	AL379190	C 486	20	1.3	409	9	AM168383	AM168383 x181f05.x
C 414	20	1.3	370	9	A1630356	C 487	20	1.3	409	10	BE304056	BE304056 UI-R-GS0-
C 415	20	1.3	370	10	BE362520	C 488	20	1.3	409	10	BE3079905	BE3079905 UI-R-CS0-
C 416	20	1.3	370	10	BE3059835	C 489	20	1.3	409	10	BE381160	BE381160 UI-R-CS0-
C 417	20	1.3	370	10	BE994680	C 490	20	1.3	409	10	BE991431	BE991431 UI-M-B21-
C 418	20	1.3	370	10	BE455356	C 491	20	1.3	410	9	A1281048	A1281048 4A3A-ABA-
C 419	20	1.3	371	9	AM122297	C 492	20	1.3	410	9	AU082987	AU082987 AU082987
C 420	20	1.3	372	9	A1633568	C 493	20	1.3	410	10	BE995402	BE995402 UI-M-CG0P
C 421	20	1.3	372	10	BM321105	C 494	20	1.3	411	9	A1848379	A1848379 UI-M-AH1-
C 422	20	1.3	373	9	AL386807	C 495	20	1.3	411	9	BE115471	BE115471 UI-R-BS1-
C 423	20	1.3	373	9	AA452663	C 496	20	1.3	411	10	BE947562	BE947562 UI-M-BH3-
C 424	20	1.3	374	10	BE408343	C 497	20	1.3	412	9	A1112029	A1112029 UI-R-YO-m
C 425	20	1.3	375	9	AM192111	C 498	20	1.3	412	10	BM344034	BM344034 UI-M-AP0-
C 426	20	1.3	375	10	BI299388	C 499	20	1.3	413	10	HO6065	HO6065 t145h08.y
C 427	20	1.3	376	10	BI295786	C 500	20	1.3	414	9	A1840473	A1840473 UI-M-ANO-
C 428	20	1.3	377	9	BE108471	C 501	20	1.3	415	9	AA761915	AA761915 n242g11.s
C 429	20	1.3	378	10	BE668404	C 502	20	1.3	415	10	BE955667	BE955667 UI-M-BH4-
C 430	20	1.3	379	9	A1853989	C 503	20	1.3	416	9	A1791933	A1791933 n166e07.y
C 431	20	1.3	379	10	BE629739	C 504	20	1.3	416	9	A1791934	A1791934 n167e07.y
C 432	20	1.3	379	10	BM145111	C 505	20	1.3	417	10	BM419332	BM419332 R020H04.T1
C 433	20	1.3	379	10	BE646953	C 506	20	1.3	417	10	RO6827	RO6827 y11904.r1
C 434	20	1.3	380	10	BE001344	C 507	20	1.3	418	9	AU164471	AU164471 AU164471
C 435	20	1.3	381	9	AA899839	C 508	20	1.3	419	9	AV564344	AV564344 AV564344
C 436	20	1.3	381	9	BE113565	C 509	20	1.3	419	10	BI303736	BI303736 UI-R-DP0-
C 437	20	1.3	381	10	BE411276	C 510	20	1.3	420	9	A1553537	A1553537 UI-R-C2P-
C 438	20	1.3	383	9	AA926019	C 511	20	1.3	420	9	BB689501	BB689501 BB689501
C 439	20	1.3	383	9	AA961140	C 512	20	1.3	421	9	A1732642	A1732642 n167e07.x
C 440	20	1.3	383	10	BE373741	C 513	20	1.3	421	9	AU041697	AU041697 AU041697
C 441	20	1.3	383	10	BE518547	C 514	20	1.3	421	10	BI277285	BI277285 UI-R-CV0-
C 442	20	1.3	384	10	BI305791	C 515	20	1.3	421	12	AQ262290	AQ262290 C17BI-E1-
C 443	20	1.3	384	10	BI306741	C 516	20	1.3	422	9	AM135551	AM135551 UI-H-B11-
C 444	20	1.3	386	9	A1556503	C 517	20	1.3	422	12	BI538626	BI538626 434038.MA
C 445	20	1.3	386	9	AA614066	C 518	20	1.3	422	12	AQ223926	AQ223926 HS-2259.B
C 446	20	1.3	386	10	BM419968	C 519	20	1.3	423	9	AA859669	AA859669 UI-R-E0-b
C 447	20	1.3	389	9	A1556375	C 520	20	1.3	423	9	AA578497	AA578497 n166g07.s
C 448	20	1.3	389	9	AA515453	C 521	20	1.3	425	9	AA960495	AA960495 vW63c09.s
C 449	20	1.3	389	10	BM373629	C 522	20	1.3	425	9	AM522144	AM522144 UI-R-B00-
C 450	20	1.3	391	9	A1219579	C 523	20	1.3	425	9	BE111918	BE111918 UI-R-BS1-
C 451	20	1.3	391	10	AM304294	C 524	20	1.3	426	9	A1384231	A1384231 fD13c11.x
C 452	20	1.3	391	10	BM383290	C 525	20	1.3	426	9	A1835212	A1835212 UI-M-AQ0-
C 453	20	1.3	391	10	W52121	C 526	20	1.3	426	12	A2278816	A2278816 RPCI-23-1
C 454	20	1.3	391	10	BE511000	C 527	20	1.3	427	9	A1566750	A1566750 UI-R-C2P-
C 455	20	1.3	392	9	A112848	C 528	20	1.3	427	9	AM449306	AM449306 UI-H-B13-

529	20	1.3	427	9	BE023288	BE023288 sm80d07.y	602	20	1.3	462	10	C98776	C98776	509	10	B1801152	B1801152
530	20	1.3	427	10	BE133963	BE133963 UI-M-BH3-	603	20	1.3	462	10	BE513153	BE513153	509	10	B1801152	B1801152
531	20	1.3	427	10	BE399275	BE399275 UI-R-CA1-	604	20	1.3	463	10	BE097186	BE097186	509	10	B1801152	B1801152
532	20	1.3	428	9	AA765994	AA765994 oad26h09.s	605	20	1.3	463	10	BM219638	BM219638	509	10	B1801152	B1801152
533	20	1.3	428	9	AI845339	AI845339 UI-M-BG0-	606	20	1.3	463	10	BE248192	BE248192	509	10	B1801152	B1801152
534	20	1.3	428	10	BI402885	BI402885 MI-P-CPI-	607	20	1.3	464	10	AI839502	AI839502	509	10	B1801152	B1801152
535	20	1.3	428	10	BE716607	BE716607 NCESTR3a15	608	20	1.3	464	10	BE373188	BE373188	509	10	B1801152	B1801152
536	20	1.3	429	9	AI847461	AI847461 UI-M-A11-	609	20	1.3	466	9	AU174900	AU174900	509	10	B1801152	B1801152
537	20	1.3	429	9	BE117204	BE117204 UI-R-B51-	610	20	1.3	466	9	BE105604	BE105604	509	10	B1801152	B1801152
538	20	1.3	429	9	BE118965	BE118965 UI-R-B51-	611	20	1.3	466	10	BE457068	BE457068	509	10	B1801152	B1801152
539	20	1.3	429	10	BE409214	BE409214 UI-R-BT1-	612	20	1.3	468	9	AM065504	AM065504	509	10	B1801152	B1801152
540	20	1.3	430	9	AA768669	AA768669 oad7912.s	613	20	1.3	469	9	AI793733	AI793733	509	10	B1801152	B1801152
541	20	1.3	430	9	AV743024	AV743024 AV473024	614	20	1.3	469	9	BB698471	BB698471	509	10	B1801152	B1801152
542	20	1.3	430	9	AM355678	AM355678 pftlc.pk0	615	20	1.3	470	9	AI840118	AI840118	509	10	B1801152	B1801152
543	20	1.3	431	9	AM874529	AM874529 hp95403.x	616	20	1.3	472	9	AI716091	AI716091	509	10	B1801152	B1801152
544	20	1.3	431	9	BE098148	BE098148 UI-R-BJ1-	617	20	1.3	472	10	BM276299	BM276299	509	10	B1801152	B1801152
545	20	1.3	431	10	BE291209	BE291209 602388655	618	20	1.3	473	9	AM254247	AM254247	509	10	B1801152	B1801152
546	20	1.3	432	9	AM158274	AM158274 2a39c08.x	619	20	1.3	474	9	AT004712	AT004712	509	10	B1801152	B1801152
547	20	1.3	433	9	AI575420	AI575420 UI-R-Y0-V	620	20	1.3	474	12	AZ061856	AZ061856	509	10	B1801152	B1801152
548	20	1.3	434	9	AI732641	AI732641 n166q07.x	621	20	1.3	475	9	AI732517	AI732517	509	10	B1801152	B1801152
549	20	1.3	434	9	AM958320	AM958320 EST370390	622	20	1.3	475	9	AI840217	AI840217	509	10	B1801152	B1801152
550	20	1.3	434	12	BA9375	BA9375 RPT111-6D18	623	20	1.3	475	9	AM049091	AM049091	509	10	B1801152	B1801152
551	20	1.3	435	9	AI850306	AI850306 UI-M-BG1-	624	20	1.3	475	9	AM049091	AM049091	509	10	B1801152	B1801152
552	20	1.3	435	9	BE096575	BE096575 UI-R-BU0-	625	20	1.3	475	9	AM049091	AM049091	509	10	B1801152	B1801152
553	20	1.3	435	10	BE093613	BE093613 UI-M-B21-	626	20	1.3	476	9	AA805508	AA805508	509	10	B1801152	B1801152
554	20	1.3	436	9	AM433868	AM433868 UI-R-BJ0P	627	20	1.3	476	9	BE098524	BE098524	509	10	B1801152	B1801152
555	20	1.3	436	9	BE102585	BE102585 UI-R-BT1-	628	20	1.3	477	10	BE448028	BE448028	509	10	B1801152	B1801152
556	20	1.3	436	10	BE1759638	BE1759638 603047082	629	20	1.3	478	10	BE93387	BE93387	509	10	B1801152	B1801152
557	20	1.3	437	9	AI313092	AI313092 fad7904.x	630	20	1.3	479	10	BE392451	BE392451	509	10	B1801152	B1801152
558	20	1.3	437	10	BI810524	BI810524 J010F08.O	631	20	1.3	480	9	AI500176	AI500176	509	10	B1801152	B1801152
559	20	1.3	438	9	AI840253	AI840253 UI-M-AH0-	632	20	1.3	480	12	AO941790	AO941790	509	10	B1801152	B1801152
560	20	1.3	438	9	AI841628	AI841628 zc97908.f	633	20	1.3	481	10	BI354378	BI354378	509	10	B1801152	B1801152
561	20	1.3	438	10	AV736142	AV736142 AV736142	634	20	1.3	483	9	AI864013	AI864013	509	10	B1801152	B1801152
562	20	1.3	438	10	BE377790	BE377790 UI-R-CU0-	635	20	1.3	483	9	AU094483	AU094483	509	10	B1801152	B1801152
563	20	1.3	438	10	CI8256	CI8256 C18256 Humna	636	20	1.3	484	10	BE110465	BE110465	509	10	B1801152	B1801152
564	20	1.3	439	9	AI201374	AI201374 qf71102.x	637	20	1.3	485	10	BE512505	BE512505	509	10	B1801152	B1801152
565	20	1.3	439	12	AA0406831	AA0406831 HS_5092.B	638	20	1.3	486	12	AZ616665	AZ616665	509	10	B1801152	B1801152
566	20	1.3	440	9	BE113876	BE113876 UI-R-CA0-	639	20	1.3	486	12	AZ616665	AZ616665	509	10	B1801152	B1801152
567	20	1.3	440	10	BI302654	BI302654 UI-R-DD0-	640	20	1.3	487	9	AI839728	AI839728	509	10	B1801152	B1801152
568	20	1.3	442	9	AI167706	AI167706 o289f10.x	641	20	1.3	487	9	AI839728	AI839728	509	10	B1801152	B1801152
569	20	1.3	442	10	BE242565	BE242565 TCAAP1D15	642	20	1.3	487	10	BI998324	BI998324	509	10	B1801152	B1801152
570	20	1.3	442	10	BE714603	BE714603 mab01H02.	643	20	1.3	488	9	AI521472	AI521472	509	10	B1801152	B1801152
571	20	1.3	443	9	AI713065	AI713065 UI-R-Y0-a	644	20	1.3	488	10	BE388529	BE388529	509	10	B1801152	B1801152
572	20	1.3	443	9	AM297806	AM297806 UI-H-BW0-	645	20	1.3	490	10	BI316068	BI316068	509	10	B1801152	B1801152
573	20	1.3	443	9	AM613525	AM613525 bh31h06.x	646	20	1.3	490	10	BM378910	BM378910	509	10	B1801152	B1801152
574	20	1.3	445	9	AI8182617	AI8182617 UI-M-BH3-	647	20	1.3	492	12	AO080162	AO080162	509	10	B1801152	B1801152
575	20	1.3	445	9	AV713159	AV713159 AV713159	648	20	1.3	493	9	AM740688	AM740688	509	10	B1801152	B1801152
576	20	1.3	446	9	AL640817	AL640817 AL640817	649	20	1.3	494	10	BI297999	BI297999	509	10	B1801152	B1801152
577	20	1.3	446	9	AM487835	AM487835 UI-M-BH3-	650	20	1.3	495	9	AI602504	AI602504	509	10	B1801152	B1801152
578	20	1.3	446	10	BE964136	BE964136 602828904	651	20	1.3	496	10	BE542674	BE542674	509	10	B1801152	B1801152
579	20	1.3	447	9	AI045361	AI045361 UI-R-C1-K	652	20	1.3	498	9	AI146046	AI146046	509	10	B1801152	B1801152
580	20	1.3	447	9	AI0172588	AI0172588 AU172588	653	20	1.3	498	10	BE712771	BE712771	509	10	B1801152	B1801152
581	20	1.3	449	9	AI835887	AI835887 UI-M-A10-	654	20	1.3	500	9	AM532067	AM532067	509	10	B1801152	B1801152
582	20	1.3	449	9	AI0184281	AI0184281 AU184281	655	20	1.3	500	10	BE416926	BE416926	509	10	B1801152	B1801152
583	20	1.3	451	9	AI060124	AI060124 UI-R-C1-L	656	20	1.3	501	9	AI836244	AI836244	509	10	B1801152	B1801152
584	20	1.3	451	9	AI753272	AI753272 c707911.x	657	20	1.3	503	9	AV565629	AV565629	509	10	B1801152	B1801152
585	20	1.3	451	10	BM080330	BM080330 MEST107-A	658	20	1.3	503	10	BE275086	BE275086	509	10	B1801152	B1801152
586	20	1.3	451	10	RA2740	RA2740 y906C04.s1	659	20	1.3	504	10	AL372177	AL372177	509	10	B1801152	B1801152
587	20	1.3	452	9	AA807990	AA807990 nu98N04.s	660	20	1.3	504	10	BE684241	BE684241	509	10	B1801152	B1801152
588	20	1.3	452	9	AI059226	AI059226 UI-R-C1-L	661	20	1.3	504	10	BE507554	BE507554	509	10	B1801152	B1801152
589	20	1.3	453	9	AV741708	AV741708 AV741708	662	20	1.3	505	9	AA073291	AA073291	509	10	B1801152	B1801152
590	20	1.3	454	9	AI687558	AI687558 hp97a10.x	663	20	1.3	505	9	AI521184	AI521184	509	10	B1801152	B1801152
591	20	1.3	454	9	AV540583	AV540583 AV540583	664	20	1.3	506	9	AV737170	AV737170	509	10	B1801152	B1801152
592	20	1.3	456	9	AI144896	AI144896 UI-R-BT0-	665	20	1.3	506	10	BE392714	BE392714	509	10	B1801152	B1801152
593	20	1.3	456	9	AM298816	AM298816 UI-H-BW0-	666	20	1.3	507	9	AI501434	AI501434	509	10	B1801152	B1801152
594	20	1.3	458	9	AI299084	AI299084 qn35h05.x	667	20	1.3	507	9	AM329803	AM329803	509	10	B1801152	B1801152
595	20	1.3	458	10	BM233942	BM233942 k043F08-	668	20	1.3	508	10	AI716615	AI716615	509	10	B1801152	B1801152
596	20	1.3	458	10	BM233642	BM233642 h04c07.x	669	20	1.3	508	10	BM280347	BM280347	509	10	B1801152	B1801152
597	20	1.3	459	9	AM134739	AM134739 UI-H-B11-	670	20	1.3	508	10	BM387476	BM387476	509	10	B1801152	B1801152
598	20	1.3	459	10	BM276216	BM276216 PFEST0a7	671	20	1.3	509	9	AI395770	AI395770	509	10	B1801152	B1801152
599	20	1.3	462	9	AI502077	AI502077 UI-R-C1-K	672	20	1.3	509	9	AL370508	AL370508	509	10	B1801152	B1801152

c 675	20	1.3	510	9	AM138907	UI-H-B11-	748	20	1.3	554	9	AF201273	AF201273
c 676	20	1.3	510	10	BE949082	UI-M-BH3-	c 749	20	1.3	554	10	AM290990	AM290990
c 677	20	1.3	512	9	AI138056	UI-R-C2P-	c 750	20	1.3	554	10	BE678064	BE678064
c 678	20	1.3	512	9	AM534385	UI-R-C4-a	c 751	20	1.3	555	9	BE097003	BE097003
c 679	20	1.3	512	9	AM539506	CO120F09-	c 752	20	1.3	557	10	BM154097	BM154097
c 680	20	1.3	512	10	BG381422	UI-R-CT0-	c 753	20	1.3	557	9	AA875069	AA875069
c 681	20	1.3	513	9	AI836243	UI-M-AP0-	c 754	20	1.3	559	10	BG63777	BG63777
c 682	20	1.3	513	9	AI836243	UI-M-AP0-	c 755	20	1.3	559	10	BI275694	BI275694
c 683	20	1.3	513	10	BG380507	UI-R-CT0-	c 756	20	1.3	560	10	BI310016	BI310016
c 684	20	1.3	515	9	AM296931	UI-H-BM0-	c 757	20	1.3	560	10	BI886843	BI886843
c 685	20	1.3	515	10	BE950659	UI-M-CP0-	c 758	20	1.3	561	10	BM336250	BM336250
c 686	20	1.3	516	9	AJ280858	AA3A-AAI-	c 759	20	1.3	562	9	AM137497	AM137497
c 687	20	1.3	516	9	AM060450	UI-M-BH1-	c 760	20	1.3	565	10	BM274784	BM274784
c 688	20	1.3	516	10	BI067754	pgf1n-CP0	c 761	20	1.3	565	10	BF507500	BF507500
c 689	20	1.3	518	10	BE466475	UI-M-CP0	c 762	20	1.3	566	9	AV533051	AV533051
c 690	20	1.3	520	9	AA761138	AA761138	c 763	20	1.3	566	10	BF117742	BF117742
c 691	20	1.3	520	9	AM270042	xv57f08.s	c 764	20	1.3	569	10	BE548173	BE548173
c 692	20	1.3	521	9	AM281944	fj59e04.x	c 765	20	1.3	569	12	A2967841	A2967841
c 693	20	1.3	521	10	BF403788	UI-R-CA1-	c 766	20	1.3	570	10	BM274795	BM274795
c 694	20	1.3	522	9	AI714505	UI-R-CA1-	c 767	20	1.3	572	12	A2659203	A2659203
c 695	20	1.3	522	10	BI161850	BI161850	c 768	20	1.3	574	10	BI721485	BI721485
c 696	20	1.3	522	10	BF196350	BF196350	c 769	20	1.3	575	9	AV715655	AV715655
c 697	20	1.3	523	9	AI172538	UI-R-C2P-	c 770	20	1.3	575	9	AA453166	AA453166
c 698	20	1.3	523	10	BG381203	UI-R-CT0-	c 771	20	1.3	576	9	AI714497	AI714497
c 699	20	1.3	524	9	AI713300	UI-R-ABI-	c 772	20	1.3	577	12	AQ410094	AQ410094
c 700	20	1.3	524	9	AV736529	AV736529	c 773	20	1.3	578	12	A2458849	A2458849
c 701	20	1.3	524	9	AM449230	AM449230	c 774	20	1.3	579	9	AI714504	AI714504
c 702	20	1.3	524	10	BM388489	UI-R-DX0-	c 775	20	1.3	581	10	BG487376	BG487376
c 703	20	1.3	526	10	BI277170	UI-R-CY0-	c 776	20	1.3	582	10	BG835077	BG835077
c 704	20	1.3	526	10	BF400655	UI-R-CA0-	c 777	20	1.3	585	9	BE111753	BE111753
c 705	20	1.3	527	10	BI067769	pgf1n-PK0	c 778	20	1.3	585	10	BI404546	BI404546
c 706	20	1.3	527	10	BM332534	MEST170-C	c 779	20	1.3	585	10	BF408909	BF408909
c 707	20	1.3	527	10	BF012490	BF012490	c 780	20	1.3	587	9	AM706930	AM706930
c 708	20	1.3	527	10	AA453166	HS_5194_A	c 781	20	1.3	587	9	AM970928	AM970928
c 709	20	1.3	528	12	AA939218	CP74H09.s	c 782	20	1.3	588	9	AV739462	AV739462
c 710	20	1.3	529	9	AI057550	UI-M-BH1-	c 783	20	1.3	589	9	AA797987	AA797987
c 711	20	1.3	529	9	AM049672	UI-M-BH1-	c 784	20	1.3	589	9	AI714516	AI714516
c 712	20	1.3	529	9	AM119986	707006F09	c 785	20	1.3	590	10	AI013558	AI013558
c 713	20	1.3	529	10	BM379950	MEST512-H	c 786	20	1.3	590	10	BG807236	BG807236
c 714	20	1.3	529	10	AI826069	wk28B10.x	c 787	20	1.3	591	10	BI428793	BI428793
c 715	20	1.3	530	9	AM444929	UI-H-B13-	c 788	20	1.3	591	10	BE511297	BE511297
c 716	20	1.3	530	10	BF404279	UI-R-CA1-	c 789	20	1.3	594	10	BI805257	BI805257
c 717	20	1.3	530	10	BE404279	UI-R-CA1-	c 790	20	1.3	594	10	BM339041	BM339041
c 718	20	1.3	530	12	BH061449	RPCT-24-3	c 791	20	1.3	598	10	BM347917	BM347917
c 719	20	1.3	531	9	AI0108602	AI0108602	c 792	20	1.3	599	10	BF420639	BF420639
c 720	20	1.3	532	9	AI112963	UI-R-C0-9	c 793	20	1.3	599	10	BG802901	BG802901
c 721	20	1.3	533	9	AV754613	AV754613	c 794	20	1.3	600	10	BG835078	BG835078
c 722	20	1.3	533	10	BE440045	HTM1-816F	c 795	20	1.3	600	10	BM382367	BM382367
c 723	20	1.3	534	9	AI834736	AI834736	c 796	20	1.3	604	9	AI982412	AI982412
c 724	20	1.3	534	10	BE946259	UI-M-BH3-	c 797	20	1.3	604	10	BE509219	BE509219
c 725	20	1.3	534	10	BE946259	UI-M-BH3-	c 798	20	1.3	604	10	BE509219	BE509219
c 726	20	1.3	534	12	AQ917303	T233284B	c 799	20	1.3	606	9	BE098809	BE098809
c 727	20	1.3	535	10	BE651817	UI-M-AP0-	c 800	20	1.3	607	9	AL585481	AL585481
c 728	20	1.3	536	9	AM135047	UI-H-B11-	c 801	20	1.3	610	9	AM205382	AM205382
c 729	20	1.3	540	10	BI344041	BI344041	c 802	20	1.3	610	10	BF398606	BF398606
c 730	20	1.3	541	9	AI841813	UI-M-ANO-	c 803	20	1.3	612	10	BE681924	BE681924
c 731	20	1.3	541	10	BE669073	DRN05C05	c 804	20	1.3	612	10	BM276497	BM276497
c 732	20	1.3	542	9	AI575029	UI-R-G0-u	c 805	20	1.3	612	12	AQ972343	AQ972343
c 733	20	1.3	542	10	BG379482	UI-R-CS0-	c 806	20	1.3	613	9	AI714498	AI714498
c 734	20	1.3	544	9	AM296798	UI-H-BM0-	c 807	20	1.3	613	10	BE087268	BE087268
c 735	20	1.3	544	10	BM384134	UI-R-DX0-	c 808	20	1.3	615	9	AL500113	AL500113
c 736	20	1.3	546	10	BE673578	DRN03G03	c 809	20	1.3	615	12	AG053783	AG053783
c 737	20	1.3	547	9	AM251158	UI-R-BJ0-	c 810	20	1.3	618	10	BG379351	BG379351
c 738	20	1.3	548	9	AL584003	AL584003	c 811	20	1.3	619	9	AM091504	AM091504
c 739	20	1.3	549	9	AV543498	AV543498	c 812	20	1.3	626	10	BM386877	BM386877
c 740	20	1.3	549	9	AM531891	UI-R-BSD-	c 813	20	1.3	632	9	AI714874	AI714874
c 741	20	1.3	549	10	BM418528	R001A02 O	c 814	20	1.3	636	9	BE108211	BE108211
c 742	20	1.3	550	10	BF393807	UI-R-CA0-	c 815	20	1.3	636	10	BE6885308	BE6885308
c 743	20	1.3	551	9	AA859508	UI-R-E0-b	c 816	20	1.3	645	9	BE110554	BE110554
c 744	20	1.3	551	9	AA201992	ID05062.5	c 817	20	1.3	646	10	BI562546	BI562546
c 745	20	1.3	552	9	AM154085	f122c07.x	c 818	20	1.3	647	10	BE680287	BE680287
c 746	20	1.3	553	10	BM274623	PIEST0a44	c 819	20	1.3	648	9	AV724415	AV724415
c 747	20	1.3	554	9	AF201193	AF201193	c 820	20	1.3	650	10	C97989	C97989

821	20	1.3	651	10	BE312846	601146107	894	20	1.3	839	10	BF540920	BF540920	602067656
822	20	1.3	652	10	BI390392	PPPIC.pko	c 895	20	1.3	840	12	A0751817	A0751817	HS_5569.B
823	20	1.3	653	10	BI312259	EST531400	896	20	1.3	841	11	AF078845	AF078845	Homo sap1
824	20	1.3	654	10	BF297032	046PBC05	897	20	1.3	845	10	BG387410	BG387410	60246176
825	20	1.3	655	9	AL642590	AL642590	898	20	1.3	845	10	BE899369	BE899369	601681271
826	20	1.3	656	9	AL036786	AL036786	899	20	1.3	847	10	BI756798	BI756798	603024473
827	20	1.3	657	10	W48389	W48389	900	20	1.3	851	10	BF267529	BF267529	HY_CEA001
828	20	1.3	657	12	BH064002	BH064002	c 901	20	1.3	851	12	BH132843	BH132843	ENT01849F
829	20	1.3	661	10	BM268717	BM268717	c 902	20	1.3	857	12	CNS01HTX	AL144886	Anopheles
830	20	1.3	663	9	BB641337	BB641337	c 903	20	1.3	857	12	CNS01HTX	AL144886	Anopheles
831	20	1.3	663	10	BF404565	BF404565	c 904	20	1.3	854	10	BG110499	BG110499	602278925
832	20	1.3	664	9	AA753540	AA753540	c 905	20	1.3	875	10	BG367233	BG367233	HVME1001
833	20	1.3	664	12	BH454698	BH454698	906	20	1.3	876	10	BF343974	BF343974	602015669
834	20	1.3	665	10	BF240191	BF240191	907	20	1.3	892	12	AC080722	AC080722	Pan trogl
835	20	1.3	667	10	BM384384	BM384384	908	20	1.3	894	10	BF179155	BF179155	602151517
836	20	1.3	671	10	BE949481	BE949481	909	20	1.3	924	10	BF123125	BF123125	601807403
837	20	1.3	673	9	BB221493	BB221493	c 910	20	1.3	939	11	BC014503	BC014503	Homo sap1
838	20	1.3	673	10	BG840789	BG840789	c 911	20	1.3	939	11	BC014503	BC014503	Homo sap1
839	20	1.3	681	10	BF724733	BF724733	c 912	20	1.3	984	12	CNS037OR	AL515407	Tetradodon
840	20	1.3	682	12	BH003388	BH003388	c 913	20	1.3	989	9	AL515407	AL515407	Tetradodon
841	20	1.3	684	12	BH002809	BH002809	c 914	20	1.3	998	10	AL515407	AL515407	Tetradodon
842	20	1.3	687	10	BG532215	BG532215	c 915	20	1.3	1034	9	AL515407	AL515407	Tetradodon
843	20	1.3	699	9	AO161189	AO161189	c 916	20	1.3	1034	9	AL515407	AL515407	Tetradodon
844	20	1.3	699	9	AO161189	AO161189	c 917	20	1.3	1111	11	BC015102	BC015102	Homo sap1
845	20	1.3	701	12	AZ647643	AZ647643	c 918	20	1.3	1121	10	BM476808	BM476808	AGENCOURT
846	20	1.3	703	10	BE966414	BE966414	c 919	20	1.3	1314	10	BF797425	BF797425	602257076
847	20	1.3	704	10	BE612987	BE612987	c 920	20	1.3	1436	10	AK004673	AK004673	Mus muscu
848	20	1.3	705	12	BH553303	BH553303	c 921	20	1.3	1721	11	BC020285	BC020285	Homo sap1
849	20	1.3	709	9	BB309095	BB309095	c 922	20	1.3	1733	11	BC012859	BC012859	Homo sap1
850	20	1.3	711	10	C97976	C97976	c 923	20	1.3	2236	11	BC009185	BC009185	Homo sap1
851	20	1.3	714	9	AL048406	AL048406	c 924	20	1.2	23	2	HS0002743	HS0002743	Homo sap1
852	20	1.3	717	12	AO888285	AO888285	c 925	20	1.2	23	2	HS0003068	HS0003068	Homo sap1
853	20	1.3	718	9	AV713508	AV713508	c 926	20	1.2	23	2	HS0003085	HS0003085	Homo sap1
854	20	1.3	722	9	BE035676	BE035676	c 927	20	1.2	23	2	HS0003164	HS0003164	Homo sap1
855	20	1.3	722	10	BI279113	BI279113	c 928	20	1.2	30	2	HS0001042	HS0001042	Homo sap1
856	20	1.3	723	10	BI838318	BI838318	c 929	20	1.2	32	2	HS0001595	HS0001595	Homo sap1
857	20	1.3	724	10	BF624715	BF624715	c 930	20	1.2	32	2	HS0001674	HS0001674	Homo sap1
858	20	1.3	725	10	BE279434	BE279434	c 931	20	1.2	33	2	HS0003086	HS0003086	Homo sap1
859	20	1.3	725	10	BE279434	BE279434	c 932	20	1.2	33	2	HS0003165	HS0003165	Homo sap1
860	20	1.3	728	10	BG273735	BG273735	c 933	20	1.2	33	9	AL048719	AL048719	Homo sap1
861	20	1.3	729	10	BG088014	BG088014	c 934	20	1.2	33	9	AL048733	AL048733	Homo sap1
862	20	1.3	729	10	BG344353	BG344353	c 935	20	1.2	34	2	HS0003069	HS0003069	Homo sap1
863	20	1.3	731	12	BH502887	BH502887	c 936	20	1.2	49	10	BG361220	BG361220	Homo sap1
864	20	1.3	732	10	BG611288	BG611288	c 937	20	1.2	51	9	AA589046	AA589046	602827044
865	20	1.3	734	9	AL048393	AL048393	c 938	20	1.2	51	2	HS0002757	HS0002757	Homo sap1
866	20	1.3	736	10	BI294856	BI294856	c 939	20	1.2	53	10	BM142044	BM142044	602827044
867	20	1.3	738	10	BM389512	BM389512	c 940	20	1.2	56	10	BI502320	BI502320	602827044
868	20	1.3	738	10	BE881237	BE881237	c 941	20	1.2	59	10	BI742219	BI742219	602827044
869	20	1.3	748	12	AO113924	AO113924	c 942	20	1.2	63	10	BG361727	BG361727	602827044
870	20	1.3	747	9	AM963627	AM963627	c 943	20	1.2	63	10	BG962258	BG962258	602827044
871	20	1.3	757	10	BI860173	BI860173	c 944	20	1.2	65	2	HS0001786	HS0001786	Homo sap1
872	20	1.3	758	12	BH248122	BH248122	c 945	20	1.2	66	9	AI348706	AI348706	Homo sap1
873	20	1.3	762	9	AL535104	AL535104	c 946	20	1.2	66	9	AA572834	AA572834	Homo sap1
874	20	1.3	763	10	BI254402	BI254402	c 947	20	1.2	66	10	BG361618	BG361618	Homo sap1
875	20	1.3	766	10	BE966692	BE966692	c 948	20	1.2	66	12	CNS02DN9	CNS02DN9	Tetradodon
876	20	1.3	770	9	AL584124	AL584124	c 949	20	1.2	67	9	AI271601	AI271601	602827044
877	20	1.3	777	10	BI078603	BI078603	c 950	20	1.2	68	10	BG113507	BG113507	602827044
878	20	1.3	778	12	AG095725	AG095725	c 951	20	1.2	70	9	AI309514	AI309514	602827044
879	20	1.3	779	10	BE568925	BE568925	c 952	20	1.2	71	10	BM026556	BM026556	Homo sap1
880	20	1.3	787	9	AA732467	AA732467	c 953	20	1.2	74	9	AI683253	AI683253	602827044
881	20	1.3	787	10	BG674589	BG674589	c 954	20	1.2	76	9	AI309824	AI309824	602827044
882	20	1.3	789	12	AZ981604	AZ981604	c 955	20	1.2	77	9	AA980014	AA980014	602827044
883	20	1.3	790	12	BH567699	BH567699	c 956	20	1.2	80	9	AA009051	AA009051	602827044
884	20	1.3	792	12	BH517529	BH517529	c 957	20	1.2	81	10	N82562	N82562	602827044
885	20	1.3	795	12	BH455256	BH455256	c 958	20	1.2	81	10	T25577	T25577	602827044
886	20	1.3	799	10	BG740298	BG740298	c 959	20	1.2	83	10	BG361896	BG361896	602827044
887	20	1.3	799	12	BH455256	BH455256	c 960	20	1.2	83	10	BG362080	BG362080	602827044
888	20	1.3	805	9	AO751130	AO751130	c 961	20	1.2	85	10	BG362352	BG362352	602827044
889	20	1.3	809	10	BI651159	BI651159	c 962	20	1.2	85	10	BG817145	BG817145	602827044
890	20	1.3	810	10	BG442801	BG442801	c 963	20	1.2	86	10	BM026438	BM026438	602827044
891	20	1.3	827	10	BF984513	BF984513	c 964	20	1.2	88	9	AA582651	AA582651	602827044
892	20	1.3	838	10	BF264841	BF264841	c 965	20	1.2	88	10	BE662018	BE662018	602827044
893	20	1.3	838	12	AZ201255	AZ201255	c 966	20	1.2	89	9	AI223740	AI223740	602827044

```

967 19 1.2 89 12 CNS04J97 AL293236 Tetradon
968 19 1.2 94 9 AJ318247 AJ318247
969 19 1.2 99 2 HSM001797 AJ1037466 Homo sap1
970 19 1.2 100 9 AI310259 AI310259
971 19 1.2 100 9 AU038797 AU038797
972 19 1.2 101 9 AI252885 AI252885
973 19 1.2 103 9 AI431004 AI431004
974 19 1.2 104 9 AI250713 AI250713
975 19 1.2 104 9 AJ318296 AJ318296
976 19 1.2 104 9 AW336521 AW336521
977 19 1.2 104 10 BI470550 BI470550
978 19 1.2 105 9 AW415490 AW415490
979 19 1.2 105 9 BG360183 BG360183
980 19 1.2 106 9 AW262956 AW262956
981 19 1.2 106 10 BG153744 BG153744
982 19 1.2 107 10 BG360957 BG360957
983 19 1.2 111 10 BM161838 BM161838
984 19 1.2 112 10 BF883795 BF883795
985 19 1.2 112 10 A2227488 A2227488
986 19 1.2 114 10 BG151216 BG151216
987 19 1.2 115 10 BF415632 BF415632
988 19 1.2 117 9 BE043260 BE043260
989 19 1.2 119 9 AW484731 AW484731
990 19 1.2 119 10 BE987817 BE987817
991 19 1.2 121 10 BG362335 BG362335
992 19 1.2 122 10 BI280842 BI280842
993 19 1.2 123 10 BI399017 BI399017
994 19 1.2 124 10 AW091178 AW091178
995 19 1.2 124 10 BG913282 BG913282
996 19 1.2 124 10 BE757520 BE757520
997 19 1.2 125 9 AA769981 AA769981
998 19 1.2 125 9 AU165528 AU165528
999 19 1.2 128 9 AI763694 AI763694
1000 19 1.2 128 9 AL506504 AL506504

```

## ALIGNMENTS

```

RESULT 1
LOCUS AI542842 448 bp mRNA linear EST 23-APR-2001
DEFINITION SD09326.5prime SD Drosophila melanogaster Schneider L2 cell culture
port2 Drosophila melanogaster cDNA clone SD09326 5 similar to
CG11822: FBan0011822 'ion channel' located on: 2L 21C5-21C5';
04/13/2001, mRNA sequence.

```

```

ACCESSION AI542842 GI:13771762
VERSION
KEYWORDS
SOURCE
ORGANISM

```

```

REFERENCE 1 (bases 1 to 448)
AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
BDGP/HMI Drosophila EST Project
Unpublished (2001)
On Mar 19, 1999 this sequence version replaced gi:4460215.
Other_ESTs: SD09326.5prime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AB003589: arm:2L [301580,604052]
estimated-cyto:21B7-21C7: 04/13/2001
Plate: SD.93 row: C column: 2
High quality sequence stop: 366
POLYA-No.
Location/Qualifiers

```

## FEATURES

```

source 1. 448
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="SD09326"
/clone_lib="SD Drosophila melanogaster Schneider L2 cell
culture port2"
/lab_host="DH5-alpha"
/note="Vector: port2; Site_1: EcoRI, site_2: XhoI, sized
fractionated cDNAs were directly ligated into port2.
Plasmid cDNA library."
BASE COUNT 115 a 140 c 116 g 77 t
ORIGIN

```

```

Query Match 18.7%; Score 288; DB 9; Length 448;
Best local similarity 99.3%; Pred. No. 1.9e-73;
Matches 438; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 14 gtacatccgaacaaagcgcgctgaacaatgaagcagactcccaagataaagcaccag 73
DB 8 gTACATCCGAACAAAGCGCGCTGAACATGACGACACTCCCAAGATAAAGCACAG 67
QY 74 ttccggtcctgagctgcacactactctgcaaatgtctaatggygagatcttctatgggc 133
DB 68 TTTCCGGTCTCGACTGCGCCTGCTGCTCAATGCTTAATGGGAGCTTTATGGGCG 127
QY 134 tgacttcgtgcccagcgccactgcccagcgccagcccaagacgccaatgtcaagccg 193
DB 128 TGACTTCGTCGCGCAGCGCCACTGCGCAGCGCCAGCCCAAGAGCCATGTCAAGGCG 187
QY 194 tggatcgctccacgcgcgcgctgttcaagaaactagacagcagcagtgtagcggttcc 253
DB 188 TGATCGCTCCACGCGCGCGCTGTTCACGAACACAGCAGCGATGTGACGCGGTTC 247
QY 254 aaggaacccccccagcagcgttccctggaatggtgtgtccactaagaagcagcagat 313
DB 248 AAGGAACCCCAAGCAAGCTGCTCTGGAGTGGTGTCTACATGACATGACAGCACT 307
QY 314 tgaacgcaagctgaccaccacactgctgctgatactccgatgagagacgaagcgcg 373
DB 308 TGAACGCAAGCTGACCACTGCTGCTGATGATTCGATGAGAGAGAGAGCGG 367
QY 374 tgtgcaacgcgtcaaatatgacaacatcaagcagatcattgaagtcagcaggtcc 433
DB 368 TGTGCAACGCTGACATATGACACATCAGCAGATGACTTGAAGTCCAGCAGGCTCT 427
QY 434 ggaaccccccaatcacactc 454
DB 428 GGACCCCAATCACACTCT 448

```

```

RESULT 2
LOCUS CNS0137Y/C 890 bp DNA linear GSS 28-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC
BACN10E18 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

```

```

ACCESSION AL103312.1 GI:5614923
VERSION
KEYWORDS
SOURCE
ORGANISM

```

```

REFERENCE 1 (bases 1 to 890)
AUTHORS Genoscope.
TITLE Direct Submissiion
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

```

```

COMMENT determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (BDGP) -

```



http://www.edgp.ebi.ac.uk . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

# FEATURES

## source

1. 890  
/organism="Drosophila melanogaster"  
/plasmid="pBelobAC11"  
/db\_xref="taxon:7227"  
/clone\_lib="DrosBAC"  
/clone="BACN10F18"  
/note="end : SP6"  
BASE COUNT 205 a 193 c 240 g 198 t 54 others  
ORIGIN

## Query Match

Best Local Similarity 10.5%; Score 162; DB 12; Length 890;  
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 568 ctcaatgagcccaagcaagcagactgcaagtgaagatcgctcttgaggctgaag 627  
DB 310 CTCAACTGGCCCGACGACAGCAGAGCTGCAAGTGAAGATCGGCTCGGCGCTGAAG 251  
QY 628 gtgcctctgcggagagcagcgagcgagagagatcccttgaccagcgactgctgt 687  
DB 250 GTGCTCTGCGCGAGAACGGACGCGAGAGAGAGTCCCTTGACCAAGAGACTGCTT 191  
QY 688 cagtcaccgagatggaatcgtgactcgagaccacttt 729  
DB 190 CAGTCACCGAGGTGGAAATGCTGACTCGGAGCCCACTTT 149

RESULT 3 361 bp mRNA linear EST 05-JUL-2001  
AM944588/c LOCUS  
DEFINITION SP09326.3prime SD Drosophila melanogaster Schneider L2 cell culture  
ACCESSION AM944588  
VERSION AM944588  
KEYWORDS  
SOURCE  
ORGANISM

fruit fly.  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 361)  
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,  
Lewis, S. and Rubin, G.M.  
BDGP/HMI Drosophila EST Project  
Unpublished (2001)  
Other ESTs: SP09326.3prime  
Contact: Stapleton, M.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: http://www.fruitfly.org/EST\_estefruitfly.berkeley.edu  
Based upon the presence of a XhoI site followed by a run of 14 or  
more T residues at the beginning of the sequence, this clone was  
polydenylated. The resulting poly-T sequence has been removed. hit  
genomic AE003589: Drosophila melanogaster genomic scaffold  
14200013386046 section 15 of 16, complete sequence.: 03/19/2001  
Plate: SP.93 row: C column: 2  
High quality sequence stop: 281.

## FEATURES

1. 361  
Location/Qualifiers  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"

/clone="SD09326"  
/clone\_lib="SD Drosophila melanogaster Schneider L2 cell  
culture pOT2"  
/lab\_host="DH5-alpha"  
/note="vector: pOT2; Site\_1: EcoRI; Site\_2: XhoI; Sized  
fractionated cDNAs were directly ligated into pOT2.  
Plasmid cDNA library."

BASE COUNT 80 a 98 c 102 g 80 t 1 others  
ORIGIN

## Query Match

Best Local Similarity 3.7%; Score 57; DB 9; Length 361;  
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1413 tgaactagtttgattcgattcattcatgatttaattatgtgtgagactataatt 1469  
DB 99 TGAAC TAGTTTGCATTTCGATTTCATGTATTAATGTGTGCGAACTATTAATT 43

RESULT 4 303 bp mRNA linear EST 19-JUL-2001  
BI285394/c LOCUS  
DEFINITION UI-R-DB0-byz-e-12-0-UI.s1 UI-R-DB0 Rattus norvegicus cDNA clone  
UI-R-DB0-byz-e-12-0-UI 3', mRNA sequence.  
ACCESSION BI285394  
VERSION BI285394.1 GI:14938986  
KEYWORDS  
SOURCE  
ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 303)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.iowa.edu

The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. The sequence tag present in the cDNA between the NotI site  
and the oligo-dT track served to verify it as a clone from the  
non-normalized bladder library cDNA Library Preparation: M.B.  
Soares Lab Clone distribution: clones will be available through  
Research Genetics (www.resgen.com)  
Seq primer: M13 Forward  
POLYA=yes.

## FEATURES

### source

1. 303  
Location/Qualifiers  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-DB0-byz-e-12-0-UI"  
/clone\_lib="UI-R-DB0"  
/dev\_stage="ADULT"  
/lab\_host="DH10B (Life Technologies)"  
/note="vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site\_1: Not I; Site\_2: Eco RI; The UI-R-DB0  
library is a non-normalized library constructed from rat  
bladder tissue. For a detailed description of the library  
from which this clone was derived, please visit our web  
site at ratest.eng.iowa.edu. The subtraction has been  
previously described in (Bonaldo, Lennon and Soares,  
Genome Research 6:791-806, 1996)

TAG\_LIB=UI-R-DB0  
TAG\_TISSUE=Bladder  
TAG\_SEQ=AGC  
BASE COUNT 83 a 53 c 63 g 104 t  
ORIGIN

Query Match 1.6%; Score 24; DB 10; Length 303;  
Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1516 tctgcttacaataaaaaa 1539  
DB 25 TCTGCTTACAAAAA 2

RESULT 5  
AI801505/c 445 bp mRNA linear EST 14-DEC-1999  
LOCUS AI801505  
DEFINITION t090c09.x1 NCI-CGAP Gas4 Homo sapiens cDNA clone IMAGE:218552 3'  
similar to contains Alu repetitive element; mRNA sequence.  
ACCESSION AI801505  
VERSION AI801505.1 GI:5366977  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 445)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
CONTACT: Robert Strausberg, Ph.D.  
Email: [cgapbsr@mail.nih.gov](mailto:cgapbsr@mail.nih.gov)  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/dbdp/image/image.html](http://www.bio.llnl.gov/dbdp/image/image.html)  
Insert Length: 2240 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 417.

FEATURES  
Source Location/Qualifiers

1..445  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:218552"  
/clone\_lib="NCI-CGAP Gas4"  
/tissue\_type="poorly differentiated adenocarcinoma with  
signet ring cell features"  
/lab\_host="DH10B"  
/note="Organ: stomach; Vector: PCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"

BASE COUNT 134 a 100 c 100 g 111 t  
ORIGIN

Query Match 1.6%; Score 24; DB 9; Length 445;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1516 tctgcttacaataaaaaa 1539  
DB 28 TCTGCTTACAAAAA 5

RESULT 6

BI399302/c 475 bp mRNA linear EST 14-AUG-2001  
LOCUS BI399302  
DEFINITION MI-P-AV1-nrp-g-11-0-UI s1 MI-P-AV1 Sus scrofa cDNA clone  
MI-P-AV1-nrp-g-11-0-UI 3', mRNA sequence.  
ACCESSION BI399302  
VERSION BI399302.1 GI:15178363  
KEYWORDS EST.  
SOURCE Sus scrofa  
ORGANISM pig.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 475)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)

JOURNAL MEDLINE  
COMMENT 97044477  
Contact: Tuglie CK  
Molecular Genetics Laboratory, Department of Animal Science  
Iowa State University  
201 Klidde Hall, Ames, IA 50011-3150, USA  
Tel: 5152944252  
Fax: 5152942401

Email: [cktuglie@iastate.edu](mailto:cktuglie@iastate.edu)  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. The sequence tag present in the cDNA between the NotI site  
and the oligo-dT track served to verify it as a clone from the  
normalized placenta library cDNA library Preparation: M.B. Soares  
Lab, University of Iowa EST sequencing: M.B. Soares Lab, University  
of Iowa Clone distribution: clones will be available through  
Research Genetics ([www.resgen.com](http://www.resgen.com)) The following repetitive  
elements were found in this cDNA sequence: 74-95,  
>AT-rich#low\_complexity  
Seq primer: M13 Forward  
POLYA=Yes.

FEATURES  
Source Location/Qualifiers

1..475  
/organism="Sus scrofa"  
/strain="crossbreed"  
/db\_xref="taxon:9623"  
/clone="MI-P-AV1-nrp-g-11-0-UI"  
/clone\_lib="MI-P-AV1"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pRT33-pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: EcoRI; The MI-P-AV1  
library is normalized library derived from the MI-P-AV1  
library, ultimately derived from placenta tissue. For a  
detailed description of the library from which this clone  
was derived, please visit our web site at  
<http://piglet.genome.iastate.edu/>. The procedure used to  
create this library has been previously described (Bonaldo  
, Lennon and Soares, Genome Research 6: 791-806, 1996)

TAG\_LIB=MI-P-AV1  
TAG\_TISSUE=Placenta  
TAG\_SEQ=ATTGG

BASE COUNT 161 a 78 c 80 g 156 t  
ORIGIN

Query Match 1.6%; Score 24; DB 10; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1516 tctgcttacaataaaaaa 1539  
DB 410 TCTGCTTACAAAAA 387

RESULT 7  
BI285223/c 520 bp mRNA linear EST 19-JUL-2001  
LOCUS BI285223

DEFINITION UI-R-DB0-byx-e-09-0-UI-s1 UI-R-DB0 Rattus norvegicus cDNA clone  
 UI-R-DB0-byx-e-09-0-UI 3', mRNA sequence.  
 ACCESSION BI285223  
 VERSION BI285223.1 GI:14938644  
 KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 520)  
 Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 JOURNAL 97044477  
 MEDLINE  
 COMMENT Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cDNA between the NotI site  
 and the oligo-dT track served to verify it as a clone from the  
 non-normalized bladder library cDNA library preparation: M.B.  
 Soares Lab Clone distribution: clones will be available through  
 Research Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLYA=yes.

FEATURES  
 source  
 Location/Qualifiers  
 1..520  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-DB0-byx-e-09-0-UI"  
 /clone\_1lb="UI-R-DB0"  
 /dev\_stage="ADULT"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: p773D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-DB0  
 library is a non-normalized library constructed from rat  
 bladder tissue. For a detailed description of the library  
 from which this clone was derived, please visit our web  
 site at [rattest.eng.uiowa.edu](http://rattest.eng.uiowa.edu). The subtraction has been  
 previously described in (Bonaldo, Lennon and Soares,  
 Genome Research 6:791-806, 1996)  
 TAG\_1lb=UI-R-DB0  
 TAG\_TISSUE=bladder  
 TAG\_SEQ=ACC"

BASE COUNT 148 a 97 c 107 g 168 t  
 ORIGIN

Query Match 1.6%; Score 24; DB 10; Length 520;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1516 tctgtctacaaataaaaaaa 1539  
 ||||||||||||||||||||  
 Db 25 TCTGCTTACAAAAA 2

RESULT 8  
 A2958949 548 bp DNA linear GSS 27-APR-2001  
 LOCUS A2958949  
 DEFINITION 2M0226M20F Mouse 10kb plasmid UNGC2M library Mus musculus genomic  
 clone UNGC2M0226M20 F, DNA sequence.  
 ACCESSION A2958949  
 VERSION A2958949.1 GI:13830176

KEYWORDS GSS.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 548)  
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.  
 and Wright,D., Weiss,R.  
 Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts  
 Unpublished (2000)  
 JOURNAL Contact: Robert B. Weiss  
 University of Utah  
 Km. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., STC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0226 row: M column: 20  
 Seq primer: CGTGTGTAACGACGGCCACT  
 Class: plasmid ends  
 High quality sequence stop: 548.

FEATURES  
 source  
 Location/Qualifiers  
 1..548  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UNG2M0226M20"  
 /clone\_1lb="Mouse 10kb plasmid UNGC2M library"  
 /sex="Female"  
 /lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
 /note="Vector: PMD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (female) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of pMD42 (g114732114|p1AF129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

BASE COUNT 188 a 130 c 119 g 111 t  
 ORIGIN

Query Match 1.6%; Score 24; DB 12; Length 548;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1516 tctgtctacaaataaaaaaa 1539  
 ||||||||||||||||||||  
 Db 93 TCTGCTTACAAAAA 116

RESULT 9  
 AK017944 1560 bp mRNA linear HTC 19-JAN-2002  
 LOCUS AK017944  
 DEFINITION Mus musculus adult male thymus cDNA, RKEN full-length enriched  
 library, clone:5830424K06: hypothetical protein, full insert  
 sequence.



```

/lab_host="XLI-Blue MRF"/SOLR"
/note="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI;
mRNA made from human bone marrow stroma. cDNA made by
oligo-dT priming. Directionally cloned. Size-selected for
average insert size >0.5 kb. Library supplied by Dr. Lidin
Jia (NHGRI)."
BASE COUNT      50 a      36 c      39 g      73 t
ORIGIN

Query Match      1.5%; Score 23; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1517 ctgcttcaaaaaaaaaaaaaa 1539
Db 39 CTGCTTCAAAAAAAAAAAAAA 17

RESULT 11
LOCUS      A1535508      347 bp      mRNA      linear      EST 18-MAR-1999
DEFINITION UI-R-C3-sy-h-10-0-UI.s1 UI-R-C3 Rattus norvegicus cDNA clone
VERSION     A1535508
KEYWORDS    A1535508.1 GI:4449643
SOURCE      EST.
ORGANISM    Norway rat.
            Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 347)
AUTHORS    Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: mscares@blue.weeg.uiowa.edu
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:
            clones will be available through Research Genetics (www.resgen.com)
            Seq primer: M13 Forward
FEATURES
    source
        1..347
            /organism="Rattus norvegicus"
            /strain="Sprague-Dawley"
            /db_xref="taxon:10116"
            /clone="UI-R-C3-sy-h-10-0-UI"
            /clone_1lb="UI-R-C3"
            /dev_stage="adult"
            /lab_host="DH10B (Life Technologies)"
            /note="Vector: pT7AD-Pac (Pharmacia) with a modified
            polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C3
            library is a subtracted library of a series, ultimately
            derived from a mixture of individually tagged normalized
            libraries from rat placenta, adult lung, brain, liver,
            kidney, heart, spleen, ovary, muscle, and 8, 12 and 18-day
            embryos, after a series of subtractions to reduce the
            representation of cDNAs from which ESTs had already been
            generated. The following serially subtracted libraries
            were generated in this process: UI-R-C3, UI-R-C2p, UI-R-C1
            , UI-R-C0, UI-R-A1, UI-R-E1. The tag is a string of 3-5
            nucleotides present between the Not I site and the
            oligo-dT track which allows identification of the library

```

```

of origin of a clone within themixture. The subtracted
library (UI-R-C3) was constructed as follows: PCR amplified
cDNA inserts from UI-R-C2p clones from which 3' ESTs had
been derived was used as a driver in a hybridization with
the UI-R-C2p library in the form of single-stranded
circles. The remaining single-stranded circles (subtracted
library) was purified by hydroxyapatite column
chromatography, converted to double-stranded circles and
electroporated into DH10B bacteria (Life Technologies) to
generate the UI-R-C3 library. This procedure has been
previously described (Bonaldo, Lennon and Soares, Genome
Research 6:791-806, 1996)"
BASE COUNT      104 a      63 c      65 g      115 t
ORIGIN

Query Match      1.5%; Score 23; DB 9; Length 347;
Best Local Similarity 100.0%; Pred. No. 3.3e+03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1517 ctgcttcaaaaaaaaaaaaaa 1539
Db 23 CTGCTTCAAAAAAAAAAAAAA 1

RESULT 12
LOCUS      A1574766      351 bp      mRNA      linear      EST 30-MAR-1999
DEFINITION UI-R-G0-ub-e-05-0-UI.s1 UI-R-G0 Rattus norvegicus cDNA clone
VERSION     A1574766
KEYWORDS    A1574766.1 GI:4544762
SOURCE      EST.
ORGANISM    Norway rat.
            Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 351)
AUTHORS    Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: mscares@blue.weeg.uiowa.edu
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. The sequence tag present in the cDNA between the NotI site
            and the oligo-dT track served to verify it as a clone from the
            normalized ganglia library cDNA library Preparation: M.B. Soares
            Lab Clone distribution: clones will be available through Research
            Genetics (www.resgen.com)
            Seq primer: M13 Forward
FEATURES
    source
        1..351
            /organism="Rattus norvegicus"
            /strain="Sprague-Dawley"
            /db_xref="taxon:10116"
            /clone="UI-R-G0-ub-e-05-0-UI"
            /clone_1lb="UI-R-G0"
            /dev_stage="adult"
            /lab_host="DH10B (Life Technologies)"
            /note="Vector: pT7AD-Pac (Pharmacia) with a modified
            polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-G0
            library is a normalized library constructed from a
            mixture of rat tissues (nodose ganglia, dorsal root

```

ganglia, and trigeminal ganglia). The tag is a string of 6 nucleotides present between the Not I site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806 '1996'.

BASE COUNT 107 a 65 c 62 g 117 t

ORIGIN

Query Match 1.5%; Score 23; DB 9; Length 351;  
Best Local Similarity 100.0%; Pred. No. 3.3e+03;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1517 ctgcttacaataaaaaa 1539

Db 23 CTGCTTACAAAAA 1

## RESULT 13

LOCUS BF703043/C

DEFINITION MI-P-E5-abn-e-09-1-UM.s1 MI-P-E5 Sus scrofa cDNA clone

ACCESSION BF703043.1

VERSION BF703043.1

KEYWORDS GI:1198451

SOURCE EST.

ORGANISM pig.

JOURNAL Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Euteria; Cetartiodactyla; Suidae; Sus.

TITLE 1 (bases 1 to 359)

REFERENCE Bonaldo, M.F., Lennon, G. and Soares, M.B.

AUTHORS Normalization and subtraction: two approaches to facilitate gene

TITLE discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

COMMENT 97044477

CONTACT: Tugale CK

MOLECULAR GENETICS LABORATORY, Department of Animal Science

Iowa State University

201 Kildee Hall, Ames, IA 50011-3150, USA

Tel: 5152944252

Fax: 5152942401

Email: cktugale@iastate.edu

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site

and the oligo-dT track served to verify it as a clone from the

non-normalized embryo at gestational day 12 library cDNA library

Preparation: RJ Woods, JA Green, RS Prather S142 Animal Science

Research Center, Department of Animal Science, University of

Missouri-Columbia, 65211 Clone distribution: clones will be

available through Research Genetics (www.resgen.com)

Seq primer: M13 Forward

POLYA=yes.

FEATURES

Location/Qualifiers

1..359

/organism="Sus scrofa"

/strain="crossbred"

/db\_xref="taxon:9823"

/clone="MI-P-E5-abn-e-09-1-UM"

/clone\_lib="MI-P-E5"

/lab\_host="DH10B (Life Technologies)"

/note="Vector: pRT3D-Pac (Pharmacia) With a modified

polylinker. Site1: Not I; Site2: EcoRI. The MI-P-E5

library is derived from embryo at gestational day 12. For

a detailed description of the library from which this

clone was derived, please visit our web site at

http://pigest.genome.iastate.edu/.

TAG\_LIB=MI-P-E5

TAG\_TISSUE=embryo at gestational day 12

TAG\_SEQ=GTGAG

BASE COUNT 97 a 63 c 99 g 100 t

## ORIGIN

Query Match 1.5%; Score 23; DB 10; Length 359;  
Best Local Similarity 100.0%; Pred. No. 3.3e+03;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1517 ctgcttacaataaaaaa 1539

Db 24 CTGCTTACAAAAA 2

## RESULT 14

LOCUS D68664/C

DEFINITION CELK134B3F Yui Kohara unpublished cDNA Caenorhabditis elegans CDNA

ACCESSION D68664

VERSION D68664.1

KEYWORDS GI:1107386

SOURCE EST.

ORGANISM Caenorhabditis elegans.

JOURNAL Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea

AUTHORS ; Rhabditidae; Peloderae; Caenorhabditis.

TITLE 1 (bases 1 to 360)

REFERENCE Kohara, Y., Mitsuki, H., Nishigaki, A., Motohashi, T., Sugimoto, A. and

Tabara, H.

Toward an expression map of the C.elegans genome

Unpublished (1994)

CONTACT: Yui Kohara

Genome Biology Lab.

National Institute of Genetics

Yata 1111, Mishima, Shizuoka 411, Japan

Tel: 81-559-81-6854

Fax: 81-559-81-6855

Email: ykohara@lab.nig.ac.jp.

Location/Qualifiers

1..360

/organism="Caenorhabditis elegans"

/strain="CB1489 him-8(e1489)"

/db\_xref="taxon:6239"

/clone="YK134B3"

/clone\_lib="Yui Kohara unpublished cDNA"

/sex="hermaphrodite, male"

/tissue\_type="whole animal"

/dev\_stage="varied"

BASE COUNT 97 a 98 c 75 g 89 t 1 others

ORIGIN

Query Match 1.5%; Score 23; DB 10; Length 360;

Best Local Similarity 100.0%; Pred. No. 3.2e+03;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 337 tgcctgctgaatccgatggag 359

Db 125 TGCTGGCTGAATCTCGATGAG 103

RESULT 15

LOCUS D63343/C

DEFINITION CELK068F8F Yui Kohara unpublished cDNA Caenorhabditis elegans CDNA

ACCESSION D63343

VERSION D63343.1

KEYWORDS GI:1104985

SOURCE EST.

ORGANISM Caenorhabditis elegans.

JOURNAL Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea

AUTHORS ; Rhabditidae; Peloderae; Caenorhabditis.

TITLE 1 (bases 1 to 360)

REFERENCE Kohara, Y., Mitsuki, H., Nishigaki, A., Motohashi, T., Sugimoto, A. and

TITLE Toward an expression map of the C.elegans genome  
JOURNAL Unpublished (1994)  
COMMENT Contact: Yuji Kohara  
Genome Biology Lab.  
National Institute of Genetics  
Yata 1111, Mishima, Shizuoka 411, Japan  
Tel: 81-559-81-6854  
Fax: 81-559-81-6855  
Email: ykohara@lab.nig.ac.jp.

FEATURES  
Source  
1. .360  
/organism="Caenorhabditis elegans"  
/strain="CB1489 him-8(e1489)"  
/db\_xref="taxon:6239"  
/clone="yk68f8"  
/clone\_id="Yuji Kohara unpublished cDNA"  
/sex="hermaphrodite, male"  
/tissue\_type="whole animal"  
/dev\_stage="varied"

BASE COUNT 94 a 95 c 74 g 89 t 8 others  
ORIGIN

Query Match 1.5%; Score 23; DB 10; Length 360;  
Best Local Similarity 100.0%; Pred. No. 3.2e+03;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 337 Tgctgctgaatcgcgatgag 359  
|||||  
Db 133 TGCTGCTGAATCTCGATGAG 111

Search completed: July 18, 2002, 10:59:54  
Job time: 2417 sec







83	20	1.3	31994	22	AAS30619	DNA encoding novel	156	19	1.2	535	22	AAI14987	Human breast cancer
84	20	1.3	31994	22	AAS28165	Genomic sequence #	157	19	1.2	539	22	AAH68914	Human cervical cancer
85	20	1.3	32035	22	AAK81008	Human immune/haema	158	19	1.2	543	22	AAH72584	Human cervical cancer
86	20	1.3	57144	22	AAK79963	Human immune/haema	159	19	1.2	553	22	ABA21014	Human nervous system
87	20	1.3	113515	22	AAI34174	Human immune system	160	19	1.2	553	22	ABA21015	Human nervous system
88	19	1.2	21	16	AAQ75644	Reverse transcript	161	19	1.2	553	22	ABA21018	Human nervous system
89	19	1.2	24	22	AAI64873	Human serine/threo	162	19	1.2	555	22	AAI21018	Human breast cancer
90	19	1.2	51	22	AAI29361	Human SNP oligonuc	163	19	1.2	555	22	AAI21018	Human breast cancer
91	19	1.2	51	22	AAI29895	Human SNP oligonuc	164	19	1.2	555	22	AAI21018	Human breast cancer
92	19	1.2	51	22	AAI31945	Human SNP oligonuc	165	19	1.2	558	22	AAI21018	Human breast cancer
93	19	1.2	89	22	AAI98685	Human excretory re	166	19	1.2	568	22	AAI21018	Human breast cancer
94	19	1.2	89	22	AAI63081	Human kidney relat	167	19	1.2	568	22	AAI21018	Human breast cancer
95	19	1.2	93	18	AAI68437	Cleavable padlock	168	19	1.2	568	22	AAI21018	Human breast cancer
96	19	1.2	148	22	AAH71776	Human cervical can	169	19	1.2	568	22	AAI21018	Human breast cancer
97	19	1.2	185	22	AAK69821	Human immune/haema	170	19	1.2	568	22	AAI21018	Human breast cancer
98	19	1.2	186	22	AAK57231	Human immune/haema	171	19	1.2	568	22	AAI21018	Human breast cancer
99	19	1.2	234	22	AAK56917	Human immune/haema	172	19	1.2	568	22	AAI21018	Human breast cancer
100	19	1.2	254	22	AAH72083	Human cervical can	173	19	1.2	580	22	AAI21018	Human breast cancer
101	19	1.2	265	22	AAI00613	Human reproductive	174	19	1.2	614	20	AAI21018	Human breast cancer
102	19	1.2	266	24	ABK16193	Human lung tumour	175	19	1.2	621	22	AAH64775	Human breast cancer
103	19	1.2	275	21	AAA42928	Human secreted exp	176	19	1.2	624	21	AAI21018	Human breast cancer
104	19	1.2	304	21	AAI19533	Human secreted pro	177	19	1.2	628	21	AAI21018	Human breast cancer
105	19	1.2	324	22	AAI34337	Human CDNA encodin	178	19	1.2	683	22	AAH98208	Human secreted pro
106	19	1.2	338	19	AAV61226	CDNA sequence of p	179	19	1.2	687	20	AAI21018	Human secreted pro
107	19	1.2	338	19	AAV58611	Prostate tumour sp	180	19	1.2	720	22	AAI21018	Human secreted pro
108	19	1.2	338	21	AAI06374	Human immunogenic	181	19	1.2	733	20	AAI21018	Human secreted pro
109	19	1.2	338	21	AAI06374	Human prostate CDN	182	19	1.2	733	21	AAI21018	Human secreted pro
110	19	1.2	338	22	AAI01033	Human prostate tum	183	19	1.2	760	20	AAI21018	Human secreted pro
111	19	1.2	338	22	AAH93490	Human prostate-spe	184	19	1.2	793	21	AAI21018	Human secreted pro
112	19	1.2	338	22	AAH84804	Human prostate-spe	185	19	1.2	805	21	AAI21018	Human secreted pro
113	19	1.2	338	22	AAH02555	Prostate tumour an	186	19	1.2	821	22	AAH71495	Human breast cancer
114	19	1.2	343	22	AAI08344	Human breast cancer	187	19	1.2	823	22	AAI19568	Human breast cancer
115	19	1.2	345	23	AAI57638	CDNA #314 encoding	188	19	1.2	863	21	AAI21018	Human breast cancer
116	19	1.2	354	22	AAK79623	Human immune/haema	189	19	1.2	865	22	AAI21018	Human breast cancer
117	19	1.2	363	22	AAI65255	Novel human polyn	190	19	1.2	865	22	AAI21018	Human breast cancer
118	19	1.2	366	22	AAI45470	Human breast cell	191	19	1.2	865	22	AAI21018	Human breast cancer
119	19	1.2	366	22	ABA65303	Human foetal liver	192	19	1.2	865	22	AAI21018	Human breast cancer
120	19	1.2	366	22	ABA32404	Probe #10870 for g	193	19	1.2	871	22	AAI21018	Human breast cancer
121	19	1.2	366	22	AAI37117	Human brain expres	194	19	1.2	878	19	AAV73447	Human breast cancer
122	19	1.2	366	22	AAK39458	Human bone marrow	195	19	1.2	880	22	AAI21018	Human breast cancer
123	19	1.2	366	22	AAI20268	Probe #10201 for g	196	19	1.2	890	22	AAI21018	Human breast cancer
124	19	1.2	366	22	AAI05972	Probe #14156 used	197	19	1.2	899	22	AAI21018	Human breast cancer
125	19	1.2	366	22	AAI05972	Probe #3963 used t	198	19	1.2	905	22	AAI21018	Human breast cancer
126	19	1.2	366	22	AAI25580	Human breast cancer	199	19	1.2	912	22	AAI21018	Human breast cancer
127	19	1.2	369	22	AAI84575	Human polynucleoti	200	19	1.2	922	22	AAI21018	Human breast cancer
128	19	1.2	380	22	AAK57454	Human immune/haema	201	19	1.2	931	21	AAI21018	Human breast cancer
129	19	1.2	383	22	AAI87502	Human polynucleoti	202	19	1.2	931	21	AAI21018	Human breast cancer
130	19	1.2	393	22	AAI65925	Novel human polyn	203	19	1.2	931	21	AAI21018	Human breast cancer
131	19	1.2	395	22	AAI86896	Human polynucleoti	204	19	1.2	931	21	AAI21018	Human breast cancer
132	19	1.2	397	22	AAI90049	Human polynucleoti	205	19	1.2	936	22	AAI21018	Human breast cancer
133	19	1.2	403	22	AAI83197	Human polynucleoti	206	19	1.2	947	20	AAI21018	Human breast cancer
134	19	1.2	404	22	AAI86429	Human polynucleoti	207	19	1.2	960	21	AAI21018	Human breast cancer
135	19	1.2	407	22	AAI83508	Human polynucleoti	208	19	1.2	978	20	AAI21018	Human breast cancer
136	19	1.2	412	22	AAI86479	Human polynucleoti	209	19	1.2	988	22	AAI21018	Human breast cancer
137	19	1.2	414	22	AAI86474	Human polynucleoti	210	19	1.2	1001	18	AAI21018	Human breast cancer
138	19	1.2	415	22	AAI86881	Human polynucleoti	211	19	1.2	1011	20	AAI21018	Human breast cancer
139	19	1.2	415	22	AAI90006	Human polynucleoti	212	19	1.2	1028	22	AAI21018	Human breast cancer
140	19	1.2	415	22	AAI91424	Human polynucleoti	213	19	1.2	1029	22	AAI21018	Human breast cancer
141	19	1.2	416	22	AAI81439	Human polynucleoti	214	19	1.2	1061	21	AAI21018	Human breast cancer
142	19	1.2	417	22	AAK92617	Human CDNA 3'-end	215	19	1.2	1074	22	AAI21018	Human breast cancer
143	19	1.2	419	22	AAI81393	Human polynucleoti	216	19	1.2	1088	22	AAI21018	Human breast cancer
144	19	1.2	437	22	AAH71012	Human cervical can	217	19	1.2	1107	22	AAI21018	Human breast cancer
145	19	1.2	440	22	AAI88046	Human polynucleoti	218	19	1.2	1116	21	AAI21018	Human breast cancer
146	19	1.2	448	22	AAI01742	Human reproductive	219	19	1.2	1116	21	AAI21018	Human breast cancer
147	19	1.2	474	22	AAI35561	Human cardiovascular	220	19	1.2	1116	21	AAI21018	Human breast cancer
148	19	1.2	495	22	AAI23840	Human breast cancer	221	19	1.2	1119	21	AAI21018	Human breast cancer
149	19	1.2	501	22	AAI56459	Human CDNA for an	222	19	1.2	1133	20	AAI21018	Human breast cancer
150	19	1.2	501	22	AAI68171	Human immune/haema	223	19	1.2	1135	20	AAI21018	Human breast cancer
151	19	1.2	515	22	AAI16737	Human breast cancer	224	19	1.2	1149	21	AAI21018	Human breast cancer
152	19	1.2	529	22	AAI27548	CDNA encoding nove	225	19	1.2	1171	20	AAI21018	Human breast cancer
153	19	1.2	530	16	AAI43686	Medium chain speci	226	19	1.2	1178	21	AAI21018	Human breast cancer
154	19	1.2	533	22	AAI09181	Human breast cancer	227	19	1.2	1202	19	AAI21018	Human breast cancer
155	19	1.2	534	22	AAK92396	Human CDNA 3'-end	228	19	1.2	1202	19	AAI21018	Human breast cancer

229	19	1.2	1222	20	AAx83994	Human EPC1 encodin	302	19	1.2	3147	23	ABL02772	Drosophila melanog
230	19	1.2	1243	21	AAc95475	Human secreted pro	303	19	1.2	3153	22	AAK69887	Human immune/haema
231	19	1.2	1255	21	AAc08382	Human secreted pro	304	19	1.2	3163	22	AAK89450	Human digestive sy
232	19	1.2	1297	19	AAV73449	Flea saliv protei	305	19	1.2	3164	22	AAK89447	Human digestive sy
233	19	1.2	1297	19	AAV73450	Flea saliv protei	306	19	1.2	3164	22	AAK89447	Human digestive sy
234	19	1.2	1302	19	AAV40485	Human secreted pro	307	19	1.2	3437	22	ABL06440	Drosophila melanog
235	19	1.2	1307	21	AAc74248	Human secreted pro	308	19	1.2	3466	19	AAc92567	Mouse T2R26 nucleo
236	19	1.2	1318	21	AAc55912	Human secreted pro	309	19	1.2	3515	22	AAV37225	Human Bap-1 DNA.
237	19	1.2	1363	21	AAc55207	Human secreted pro	310	19	1.2	3577	23	AAc67293	Arbidiopsis thaila
238	19	1.2	1376	21	AAc23399	Human secreted pro	311	19	1.2	3589	22	AAc60891	DNA encoding novel
239	19	1.2	1392	20	AAx04395	Human secreted pro	312	19	1.2	3774	21	AAc78005	Human polynucleoti
240	19	1.2	1417	21	AAc74398	Human secreted pro	313	19	1.2	3853	21	AAc77598	Human cancer assoc
241	19	1.2	1516	22	AAc56392	ABC transport rela	314	19	1.2	3859	23	AAc78584	Human ORFX ORF3153
242	19	1.2	1542	20	AAx16151	Human Sox1 encodin	315	19	1.2	3907	23	ABL26630	DNA encoding novel
243	19	1.2	1578	19	AAV71117	CDNA encoding a hu	316	19	1.2	4007	22	AAc67078	Drosophila melanog
244	19	1.2	1578	19	AAV62388	Human 7-transmembr	317	19	1.2	4017	22	AAc72940	Human immune/haema
245	19	1.2	1584	21	AAc00015	Scorpion protein d	318	19	1.2	4058	22	AAc42016	Human immune/haema
246	19	1.2	1592	23	AAx07430	Homo sapiens secre	319	19	1.2	4091	22	AAc42016	Genomic sequence #
247	19	1.2	1599	23	AAc590606	DNA encoding novel	320	19	1.2	4189	23	AAc69085	Human small cell 1
248	19	1.2	1605	22	AAc94109	Human full-length	321	19	1.2	4196	23	AAc40488	DNA encoding novel
249	19	1.2	1701	17	AAc30372	MORF-1 cDNA. Homo	322	19	1.2	4199	20	AAc34304	Human SUG1 DNA.
250	19	1.2	1701	18	AAc61397	MORF-1 coding sequ	323	19	1.2	4199	20	AAc58296	Human PRO286 nucle
251	19	1.2	1701	21	AAc44745	Human FADD DNA. H	324	19	1.2	4199	20	AAc58296	Human TOLL protein
252	19	1.2	1736	21	AAc79724	Human secreted pro	325	19	1.2	4199	21	AAc78584	Human PRO286 nucle
253	19	1.2	1738	21	AAc59662	Human secreted pro	326	19	1.2	4203	20	AAx07074	Human KDS1 serine/
254	19	1.2	1821	21	AAc59421	Human secreted pro	327	19	1.2	4335	23	ABL15141	Drosophila melanog
255	19	1.2	1835	21	AAc53789	Murine incl common	328	19	1.2	4343	21	AAc77400	Human ORFX ORF2955
256	19	1.2	1857	21	AAc59142	Human secreted pro	329	19	1.2	4354	23	ABL20674	Drosophila melanog
257	19	1.2	1860	21	AAc62425	Human secreted pro	330	19	1.2	4720	22	AAc24153	Human secreted pro
258	19	1.2	1860	21	AAc62425	Human brain-associ	331	19	1.2	4733	22	ABA07211	Human pancreatic c
259	19	1.2	1901	21	AAc49780	DNA encoding a hyd	332	19	1.2	4733	22	AAc89843	Human digestive sy
260	19	1.2	1916	21	AAc78064	Tomato hydropoxo	333	19	1.2	4814	23	ABL02008	Human digestive sy
261	19	1.2	1916	22	AAH33274	Human cancer assoc	334	19	1.2	4874	20	AAc60103	Arbidiopsis thaila
262	19	1.2	1917	22	AAc05599	Human colon cancer	335	19	1.2	5026	22	AAc46517	Tumour suppressor
263	19	1.2	1933	10	AAc90120	Human secreted pro	336	19	1.2	5044	23	AAc81832	DNA encoding novel
264	19	1.2	1943	22	AAc07717	Human lipocortin D	337	19	1.2	5234	24	ABL32945	Human immune syste
265	19	1.2	1966	22	AAc159403	Human secreted pro	338	19	1.2	5278	22	AAc46376	Tumour suppressor
266	19	1.2	1966	22	AAc07747	Human polynucleoti	339	19	1.2	5278	24	ABL32823	Human immune syste
267	19	1.2	1966	22	AAc59700	Human secreted pro	340	19	1.2	5308	24	ABL32560	Human immune syste
268	19	1.2	2054	22	AAc03894	Human secreted pro	341	19	1.2	5407	24	ABL34091	Human immune syste
269	19	1.2	2060	22	AAc20949	Soybean arginyl-CK	342	19	1.2	5413	22	AAc46694	Tumour suppressor
270	19	1.2	2071	22	AAH33230	Human colon cancer	343	19	1.2	5470	23	ABL03536	Drosophila melanog
271	19	1.2	2192	23	ABL03537	Drosophila melanog	344	19	1.2	5518	24	ABL32216	Human immune syste
272	19	1.2	2215	22	AAH45148	Human protein kina	345	19	1.2	5546	24	ABL32390	Human immune syste
273	19	1.2	2242	22	AAc94899	Human full-length	346	19	1.2	5688	22	AAc89950	Human bone marrow
274	19	1.2	2289	24	AAc94899	Arbidiopsis cDNA e	347	19	1.2	5763	24	ABL33288	Human immune syste
275	19	1.2	2288	24	AAc24790	Human secreted pro	348	19	1.2	5922	24	ABL32450	Human immune syste
276	19	1.2	2363	22	AAc74284	Common ice plant c	349	19	1.2	5925	24	ABL33576	Human immune syste
277	19	1.2	2376	21	AAc75837	Human ORFX ORF1392	350	19	1.2	5931	22	AAc46703	Tumour suppressor
278	19	1.2	2476	22	AAH34123	Human colon cancer	351	19	1.2	5937	24	AAc61211	Human gene regulat
279	19	1.2	2484	22	AAH59146	Human polynucleoti	352	19	1.2	6002	24	ABL32953	Human immune syste
280	19	1.2	2507	21	AAc06561	Human immunogenic	353	19	1.2	6012	24	ABL34565	Human metastasis a
281	19	1.2	2507	22	AAc63770	Human prostate cDN	354	19	1.2	6050	22	AAc46791	Tumour suppressor
282	19	1.2	2507	22	AAc93677	Human prostate-spe	355	19	1.2	6050	24	ABL34128	Human immune syste
283	19	1.2	2507	22	AAc84491	Human prostate-spe	356	19	1.2	6077	24	ABL33246	Human immune syste
284	19	1.2	2507	22	AAc02742	Prostate tumour an	357	19	1.2	6093	24	ABL34170	Human immune syste
285	19	1.2	2610	24	AAc956324	Arbidiopsis cDNA e	358	19	1.2	6098	24	ABL32519	Human immune syste
286	19	1.2	2651	23	AAc87521	Petunia hybrida zp	359	19	1.2	6109	24	ABL32326	Human immune syste
287	19	1.2	2651	23	AAc64607	DNA encoding novel	360	19	1.2	6109	24	AAc61077	Human immune syste
288	19	1.2	2652	23	AAc23292	Drosophila melanog	361	19	1.2	6160	24	ABL33533	Human gene regulat
289	19	1.2	2653	23	AAc69572	Human immune/haema	362	19	1.2	6174	24	ABL33509	Human immune syste
290	19	1.2	2663	23	ABL18654	Human immune/haema	363	19	1.2	6189	22	AAc46599	Tumour suppressor
291	19	1.2	2712	22	AAH75376	P. hybrida pollen-	364	19	1.2	6270	22	AAc45492	Human metastasis a
292	19	1.2	2712	22	AAH45742	Petunia zinc fling	365	19	1.2	6270	22	ABL32405	Chemically pretrea
293	19	1.2	2738	22	AAc06007	Human neuronal apo	366	19	1.2	6270	24	ABL32405	Human immune syste
294	19	1.2	2743	24	AAc966335	Arbidiopsis cDNA e	367	19	1.2	6283	24	AAc61351	Human immune syste
295	19	1.2	2817	21	AAc53790	Murine P37ING1 cod	368	19	1.2	6412	24	AAc61145	Human gene regulat
296	19	1.2	2962	20	AAc28161	Rat Acid sensitive	369	19	1.2	6465	24	ABL32985	Human gene regulat
297	19	1.2	3027	23	ABL20562	Drosophila melanog	370	19	1.2	6558	22	AAc46549	Tumour suppressor
298	19	1.2	3038	23	ABL03730	Drosophila melanog	371	19	1.2	6577	22	ABL33357	Tumour suppressor
299	19	1.2	3048	22	AAc69888	Human immune/haema	372	19	1.2	6577	24	AAc61222	Human immune syste
300	19	1.2	3050	22	AAc69889	Human immune/haema	373	19	1.2	6591	24	ABL33251	Human immune syste
301	19	1.2	3101	11	AAc02047	Sequence encoding	374	19	1.2	6627	24	ABL33575	Human immune syste

C 375	19	1.2	6823	24	ABL33145	Human immune syste	C 448	19	1.2	18817	24	ABL34495	Human metastasis a
C 376	19	1.2	6853	24	ABL32311	Human immune syste	C 449	19	1.2	19082	24	ABL32626	Human immune syste
C 377	19	1.2	6915	24	ABL32406	Human immune syste	C 450	19	1.2	19131	22	AA546716	Tumour suppressor
C 378	19	1.2	7010	22	ABL03343	Human reproductive	C 451	19	1.2	19972	24	ABL32106	Human immune syste
C 379	19	1.2	7273	23	ABL06708	Drosophila melanog	C 452	19	1.2	21446	22	AAK70163	Human immune/haema
C 380	19	1.2	7241	24	AA561395	Human gene regulat	C 453	19	1.2	21446	22	AAK82629	Human immune/haema
C 381	19	1.2	7449	24	ABL32277	Human immune syste	C 454	19	1.2	22408	23	ABL05054	Drosophila melanog
C 382	19	1.2	7657	22	AA545477	Chemically pretrea	C 455	19	1.2	22874	23	ABL04654	Drosophila melanog
C 383	19	1.2	7657	22	AA545478	Chemically pretrea	C 456	19	1.2	22967	23	ABL12942	Drosophila melanog
C 384	19	1.2	7657	24	ABL34022	Human immune syste	C 457	19	1.2	23203	23	ABL17834	Drosophila melanog
C 385	19	1.2	7657	24	ABL34023	Human immune syste	C 458	19	1.2	23683	24	ABL34623	Human metastasis a
C 386	19	1.2	7852	22	AA542017	Genomic sequence #	C 459	19	1.2	23899	23	ABL10362	Drosophila melanog
C 387	19	1.2	8030	24	ABL33333	Human immune syste	C 460	19	1.2	37973	24	ABL34196	Human immune syste
C 388	19	1.2	8078	22	AA546492	Tumour suppressor	C 461	19	1.2	39353	22	AAK65396	Human immune/haema
C 389	19	1.2	8136	22	AA531510	Human DNA for a no	C 462	19	1.2	39353	22	AAK67232	Human immune/haema
C 390	19	1.2	8197	24	ABL34515	Human metastasis a	C 463	19	1.2	39358	22	AAK65397	Human immune/haema
C 391	19	1.2	8246	24	ABL33202	Human immune syste	C 464	19	1.2	39358	22	AAK67233	Human immune/haema
C 392	19	1.2	8277	24	ABL33328	Human immune syste	C 465	19	1.2	43795	21	AAK292583	Human immune/haema
C 393	19	1.2	8297	24	ABL34074	Human immune syste	C 466	19	1.2	48037	22	AAK84729	Human immune/haema
C 394	19	1.2	8499	23	ABL16018	Human immune syste	C 467	19	1.2	48037	22	AAK85983	Human immune/haema
C 395	19	1.2	8513	22	AA545355	Drosophila melanog	C 468	19	1.2	49380	23	ABL11838	Drosophila melanog
C 396	19	1.2	8513	22	AA545355	Chemically pretrea	C 469	19	1.2	49380	20	AAZ23904	Human LOBO homolog
C 397	19	1.2	8676	24	AA561414	Tumour suppressor	C 470	19	1.2	52691	22	AAK84435	Human immune/haema
C 398	19	1.2	8771	24	ABL33824	Human gene regulat	C 471	19	1.2	68840	20	AAK57351	Human immune/haema
C 399	19	1.2	8771	24	ABL33825	Human immune syste	C 472	19	1.2	110000	22	AAK84800	Human chromosome 6
C 400	19	1.2	8876	24	ABL34076	Human immune syste	C 473	19	1.2	113515	24	ABL34175	Human immune syste
C 401	19	1.2	8991	24	AA561295	Human gene regulat	C 474	19	1.2	144460	21	AAZ93815	Olfactory receptor
C 402	19	1.2	9122	22	AA56088	Human musculoskele	C 475	19	1.2	465237	24	ABA90193	Human oestrogen re
C 403	19	1.2	9219	22	AA546807	Tumour suppressor	C 476	19	1.2	1503900	22	AAK95240	Human neuregulin-1
C 404	19	1.2	9219	24	ABL34426	Human immune syste	C 477	19	1.2	1503900	22	AAK96733	Human neuregulin-1
C 405	19	1.2	9375	22	AAK37084	Human musculoskele	C 478	18	1.2	20	16	AAK75568	Reverse transcript
C 406	19	1.2	9375	22	AAK84379	Human immune/haema	C 479	18	1.2	20	16	AAK83959	BAVE2 gene fragmen
C 407	19	1.2	9502	22	AA546731	Tumour suppressor	C 480	18	1.2	21	16	AAO75643	Reverse transcript
C 408	19	1.2	9577	22	ABL23106	Drosophila melanog	C 481	18	1.2	21	16	AAO75645	Reverse transcript
C 409	19	1.2	9741	24	ABL33323	Human immune syste	C 482	18	1.2	21	16	AAO75646	Reverse transcript
C 410	19	1.2	10034	24	ABL33885	Human immune syste	C 483	18	1.2	24	24	AB185028	Capture oligonucle
C 411	19	1.2	10189	24	ABL34161	Human immune syste	C 484	18	1.2	24	24	AB185029	Capture oligonucle
C 412	19	1.2	10191	24	ABL34255	Human immune syste	C 485	18	1.2	25	21	AAK96251	H1A DPA1 gene PCR
C 413	19	1.2	10647	24	AA561396	Human gene regulat	C 486	18	1.2	29	18	AAK69832	Probe for tandem s
C 414	19	1.2	10717	24	ABL33694	Human immune syste	C 487	18	1.2	30	17	AAK24223	Oligonucleotide pr
C 415	19	1.2	11131	24	ABL32753	Human immune syste	C 488	18	1.2	30	17	AAK24222	Oligonucleotide pr
C 416	19	1.2	11398	24	ABL34438	Human immune syste	C 489	18	1.2	30	17	AAK24224	Oligonucleotide pr
C 417	19	1.2	11836	22	AA545394	Chemically pretrea	C 490	18	1.2	50	22	AAK31421	Human SNP oligonuc
C 418	19	1.2	11996	24	ABL34493	Human metastasis a	C 491	18	1.2	50	22	AAK31422	Human SNP oligonuc
C 419	19	1.2	12138	24	ABL33943	Human immune syste	C 492	18	1.2	50	22	AAK31423	Human SNP oligonuc
C 420	19	1.2	12248	23	ABL14532	Drosophila melanog	C 493	18	1.2	50	22	AAK31428	Human SNP oligonuc
C 421	19	1.2	12482	21	AA550353	Human Goodpasture	C 494	18	1.2	51	22	AAK31661	Human SNP oligonuc
C 422	19	1.2	12492	23	ABL06876	Drosophila melanog	C 495	18	1.2	51	22	AAK33366	Human SNP oligonuc
C 423	19	1.2	12730	19	AAV44233	Lettuce resistance	C 496	18	1.2	51	22	AAK33398	Human SNP oligonuc
C 424	19	1.2	12806	22	ABL24266	Drosophila melanog	C 497	18	1.2	52	21	AAK15328	Human secreted pro
C 425	19	1.2	13214	22	ABA18377	Human nervous syst	C 498	18	1.2	70	17	AAK17035	Human mitochondria
C 426	19	1.2	13333	22	AAK79541	Human immune/haema	C 499	18	1.2	71	21	AAK29601	Human secreted pro
C 427	19	1.2	13712	24	ABL33531	Human immune syste	C 500	18	1.2	71	20	AAK80775	Recombinant RNA ca
C 428	19	1.2	14169	22	AAK72936	Human immune/haema	C 501	18	1.2	72	22	AAK66829	Human immune/haema
C 429	19	1.2	14307	24	ABL32728	Human immune syste	C 502	18	1.2	72	22	AAK80857	Human immune/haema
C 430	19	1.2	14686	23	ABL15140	Drosophila melanog	C 503	18	1.2	72	22	AAK83532	Human immune/haema
C 431	19	1.2	14925	22	AAK03282	Human reproductive	C 504	18	1.2	72	22	AAK86588	Human immune/haema
C 432	19	1.2	14925	22	AAK04552	Human reproductive	C 505	18	1.2	78	21	AAK14386	Human secreted pro
C 433	19	1.2	15051	23	ABL03732	Drosophila melanog	C 506	18	1.2	88	22	AAI98803	Human excretory re
C 434	19	1.2	15306	22	AA545454	Chemically pretrea	C 507	18	1.2	88	22	AAI63199	Human kidney relat
C 435	19	1.2	15548	24	ABL34155	Human immune syste	C 508	18	1.2	94	18	AAV00411	3' fragment of c10
C 436	19	1.2	15833	24	ABL03060	Drosophila melanog	C 509	18	1.2	95	22	AAK04913	Human gene signatu
C 437	19	1.2	16217	24	ABL34157	Human immune syste	C 510	18	1.2	112	16	AAK25481	Human gene signatu
C 438	19	1.2	16766	24	ABL34264	Human immune syste	C 511	18	1.2	117	21	AAK16447	Human prostate can
C 439	19	1.2	16842	22	AA546411	Human immune syste	C 512	18	1.2	117	22	AAK68873	Human immune/haema
C 440	19	1.2	16842	24	AA561335	Tumour suppressor	C 513	18	1.2	117	22	AAK76691	Human immune/haema
C 441	19	1.2	16891	24	ABL32653	Human gene regulat	C 514	18	1.2	117	22	AAK83853	Human immune/haema
C 442	19	1.2	16891	24	ABL34505	Human immune syste	C 515	18	1.2	125	22	AAK81138	Human immune/haema
C 443	19	1.2	17137	24	ABL34191	Human metastasis a	C 516	18	1.2	137	22	AAK15347	Human breast cance
C 444	19	1.2	17389	24	ABL33415	Human immune syste	C 517	18	1.2	137	22	AAK24193	Human breast cance
C 445	19	1.2	17738	24	ABL33538	Human immune syste	C 518	18	1.2	139	22	ABA17705	Human nervous syst
C 446	19	1.2	18296	23	ABL03234	Drosophila melanog	C 519	18	1.2	139	22	AAK36331	Human musculoskele
C 447	19	1.2	18512	24	ABL32976	Human immune syste	C 520	18	1.2	141	22	ABA15776	Human nervous syst

C 521	18	1.2	141	22	AAK83995	Human immune/haema	C 594	18	1.2	320	22	AAU13401	Human breast cancer
C 522	18	1.2	141	22	AAK87112	Human immune/haema	C 595	18	1.2	321	24	AA561485	Lung small cell ca
C 523	18	1.2	142	22	AAU15315	Human breast cancer	C 596	18	1.2	328	22	AA186509	Human polynucleoti
C 524	18	1.2	142	22	AAU24161	Human breast cancer	C 597	18	1.2	331	22	AAK54664	Human haematologic
C 525	18	1.2	142	22	AAK70077	Human immune/haema	C 598	18	1.2	331	22	AAK64191	Human immune/haema
C 526	18	1.2	143	22	AAK65842	Human immune/haema	C 599	18	1.2	333	22	AAU14965	Human breast cancer
C 527	18	1.2	143	22	AAK90707	Human digestive sy	C 600	18	1.2	336	22	AAK69254	Human immune/haema
C 528	18	1.2	146	22	AAK87315	Human immune/haema	C 601	18	1.2	338	22	AAK35462	Human cardiovascular
C 529	18	1.2	148	22	AAK32635	Human genomic DNA	C 602	18	1.2	339	22	AAU35334	Human musculoskele
C 530	18	1.2	159	22	AAK75654	Human immune/haema	C 603	18	1.2	342	22	AA187251	Human polynucleoti
C 531	18	1.2	160	18	AAU61985	Detection probe fo	C 604	18	1.2	342	22	AA188836	Human polynucleoti
C 532	18	1.2	160	22	AAK40236	DNA encoding human	C 605	18	1.2	344	22	AA187135	Human polynucleoti
C 533	18	1.2	160	22	AAU02154	Human reproductive	C 606	18	1.2	344	24	AA596303	Arabidopsis cDNA e
C 534	18	1.2	164	22	AA199160	Human excretory re	C 607	18	1.2	349	22	AA180809	Human polynucleoti
C 535	18	1.2	164	22	AA163510	Human kidney relat	C 608	18	1.2	354	21	AAK63882	Nucleotide sequenc
C 536	18	1.2	165	21	AAU14177	Human secreted pro	C 609	18	1.2	356	22	ABA17877	Human nervous syst
C 537	18	1.2	165	22	ABA15465	Human nervous syst	C 610	18	1.2	359	22	AA188709	Human polynucleoti
C 538	18	1.2	165	22	AAK64234	Human immune/haema	C 611	18	1.2	361	22	AA184186	Human polynucleoti
C 539	18	1.2	170	20	AAH85961	Human single nucle	C 612	18	1.2	363	22	AA532910	Human genomic DNA
C 540	18	1.2	172	22	AAK81214	Human immune/haema	C 613	18	1.2	367	22	AA186612	Human breast cancer
C 541	18	1.2	177	22	AAK65967	Human immune/haema	C 614	18	1.2	369	22	AA190989	Human polynucleoti
C 542	18	1.2	180	21	AAU55658	Human differential	C 615	18	1.2	371	14	AAU061288	Human polynucleoti
C 543	18	1.2	181	20	AAU13491	Human gene express	C 616	18	1.2	371	23	AA585359	Human brain Expres
C 544	18	1.2	192	21	AAU12149	Human secreted pro	C 617	18	1.2	375	22	AA187531	Human polynucleoti
C 545	18	1.2	192	22	AAK77059	Human immune/haema	C 618	18	1.2	375	22	AA190980	Human polynucleoti
C 546	18	1.2	192	22	AAK7060	Human immune/haema	C 619	18	1.2	378	22	AA181567	Human polynucleoti
C 547	18	1.2	211	22	AAK67607	Human immune/haema	C 620	18	1.2	380	22	AA182817	Human polynucleoti
C 548	18	1.2	217	22	AAK32877	Human ovarian PCR-	C 621	18	1.2	380	22	AA186586	Human polynucleoti
C 549	18	1.2	217	22	AAH82434	Human ovarian tumo	C 622	18	1.2	380	22	AAK55792	Human immune/haema
C 550	18	1.2	218	22	AA524704	Human ovarian PCR-	C 623	18	1.2	381	22	AA188668	Human polynucleoti
C 551	18	1.2	218	22	AAH83332	Human ovarian tumo	C 624	18	1.2	382	22	AA187149	Human polynucleoti
C 552	18	1.2	220	22	AAK65830	Human immune/haema	C 625	18	1.2	382	22	AA187691	Human polynucleoti
C 553	18	1.2	224	19	AAU05742	Nucleotide sequenc	C 626	18	1.2	385	22	AA193254	Human polynucleoti
C 554	18	1.2	224	22	AA524213	Human ovarian PCR-	C 627	18	1.2	387	21	AAZ94888	Rat taste specific
C 555	18	1.2	224	22	AAH82794	Human ovarian tumo	C 628	18	1.2	389	22	AA187865	Human polynucleoti
C 556	18	1.2	226	22	AAU34674	Human musculoskele	C 629	18	1.2	389	22	AA188391	Human polynucleoti
C 557	18	1.2	231	21	AAZ80203	Human colon cancer	C 630	18	1.2	389	22	AAK55411	Human immune/haema
C 558	18	1.2	232	22	AAH87963	Peppermint plant o	C 631	18	1.2	390	22	AAU09940	Human breast cancer
C 559	18	1.2	240	22	AA524341	Human ovarian PCR-	C 632	18	1.2	390	22	AAU17484	Human breast cancer
C 560	18	1.2	241	22	AAH82930	Human ovarian tumo	C 633	18	1.2	391	22	AA183168	Human polynucleoti
C 561	18	1.2	242	22	AA524979	Human ovarian PCR-	C 634	18	1.2	392	22	AA186653	Human polynucleoti
C 562	18	1.2	242	22	AAH83624	Human ovarian tumo	C 635	18	1.2	395	22	AAH35297	Human colon cancer
C 563	18	1.2	247	21	AAU93738	Cat flea hindgut a	C 636	18	1.2	397	22	AA185922	Human polynucleoti
C 564	18	1.2	247	21	AAU94506	Cat flea hindgut a	C 637	18	1.2	398	22	AA183988	Human polynucleoti
C 565	18	1.2	249	22	AAU10931	Human breast cancer	C 638	18	1.2	398	22	AAK73447	Human immune/haema
C 566	18	1.2	249	22	AAU18695	Human breast cancer	C 639	18	1.2	399	22	AA185422	Human polynucleoti
C 567	18	1.2	249	22	AAK79399	Human immune/haema	C 640	18	1.2	399	22	AA187654	Human polynucleoti
C 568	18	1.2	251	22	AAU10961	Human breast cancer	C 641	18	1.2	400	21	AAK93683	Cat flea hindgut a
C 569	18	1.2	253	22	AAU198673	Human excretory re	C 642	18	1.2	401	22	AAK96165	Human neuregulin g
C 570	18	1.2	253	22	AA527556	CDNA encoding nove	C 643	18	1.2	401	22	AAK96166	Human neuregulin g
C 571	18	1.2	253	22	AA163069	Human kidney relat	C 644	18	1.2	401	22	AAK96200	Human neuregulin g
C 572	18	1.2	258	20	AAU09018	Differentiation in	C 645	18	1.2	401	22	AAK96708	Human neuregulin g
C 573	18	1.2	258	22	AAK65732	Human immune/haema	C 646	18	1.2	401	22	AAK97658	Human neuregulin g
C 574	18	1.2	270	22	AAK67467	Human immune/haema	C 647	18	1.2	401	22	AAK97659	Human neuregulin g
C 575	18	1.2	274	21	AAU00718	Human colon cancer	C 648	18	1.2	401	22	AAK97693	Human neuregulin g
C 576	18	1.2	276	20	AAV89725	EST clone CT706.	C 649	18	1.2	401	22	AAK98201	Human neuregulin g
C 577	18	1.2	279	22	AA524357	Human ovarian PCR-	C 650	18	1.2	401	22	AA192107	Human polynucleoti
C 578	18	1.2	279	22	AAH82947	Human ovarian tumo	C 651	18	1.2	403	22	AAK96719	Human polynucleoti
C 579	18	1.2	281	22	AAU28986	Colon tumour relat	C 652	18	1.2	403	22	AAK98212	Human neuregulin g
C 580	18	1.2	284	22	AA529083	CDNA encoding for	C 653	18	1.2	403	22	AAU12000	Human breast cancer
C 581	18	1.2	290	22	AAU11318	Human breast cancer	C 654	18	1.2	403	22	AAU182265	Human polynucleoti
C 582	18	1.2	294	22	AAK66970	Human immune/haema	C 655	18	1.2	404	22	AAU09899	Human breast cancer
C 583	18	1.2	295	22	AAU23818	Human breast cancer	C 656	18	1.2	404	22	AAU19697	Human breast cancer
C 584	18	1.2	296	14	AAU059461	Human brain Expres	C 657	18	1.2	405	22	AA188744	Human polynucleoti
C 585	18	1.2	300	22	AA556393	ABC transport rela	C 658	18	1.2	407	22	AAK96662	Human neuregulin g
C 586	18	1.2	302	22	AAK63052	Human immune/haema	C 659	18	1.2	407	22	AAK98135	Human neuregulin g
C 587	18	1.2	304	14	AAU05952	Human brain Expres	C 660	18	1.2	407	22	AAU15488	Human breast cancer
C 588	18	1.2	304	22	AAU05753	Human reproductive	C 661	18	1.2	409	22	AA189271	Human polynucleoti
C 589	18	1.2	304	22	AAK84486	Human immune/haema	C 662	18	1.2	412	22	AA185069	Human polynucleoti
C 590	18	1.2	305	22	AAK63278	Human immune/haema	C 663	18	1.2	412	22	AA186530	Human polynucleoti
C 591	18	1.2	306	22	AAU20217	Human breast cancer	C 664	18	1.2	413	22	AAK65941	Human immune/haema
C 592	18	1.2	308	22	AAK70443	Human immune/haema	C 665	18	1.2	415	22	AA188281	Human polynucleoti
C 593	18	1.2	314	22	AAU00786	Human reproductive	C 666	18	1.2	419	21	AAU59229	Human secreted pro

667	18	1.2	419	21	AAC69548	Human secreted pro
668	18	1.2	419	22	AA180417	Human polynucleoti
669	18	1.2	419	22	AAS02417	Human secreted pro
670	18	1.2	420	22	AA110117	Human breast cancer
671	18	1.2	421	21	AA280047	Human colon cancer
672	18	1.2	423	22	AA185265	Human polynucleoti
673	18	1.2	425	21	AA437895	EST sequence #7.
674	18	1.2	425	22	AA186980	Human polynucleoti
675	18	1.2	426	24	AAS96331	Arabidopsis CDNA e
676	18	1.2	427	22	AA193379	Human polynucleoti
677	18	1.2	431	22	AA187912	Human polynucleoti
678	18	1.2	434	22	AA124333	Human breast cancer
679	18	1.2	437	22	AAK87407	Human immune/haema
680	18	1.2	438	20	AAZ06770	PDE homologous seq
681	18	1.2	439	22	AAK55923	Human immune/haema
682	18	1.2	442	22	AA185193	Human polynucleoti
683	18	1.2	443	22	ABA57053	Human foetal liver
684	18	1.2	443	22	AAK88750	Human digestive sy
685	18	1.2	443	22	AA531784	Human liver associ
686	18	1.2	446	22	AA137502	Human musculoskele
687	18	1.2	446	22	AA137503	Human musculoskele
688	18	1.2	448	22	AA122802	Human breast cancer
689	18	1.2	450	22	AA117849	Human breast cancer
690	18	1.2	452	22	AA102233	Human reproductive
691	18	1.2	452	22	AA162531	Human breast or ov
692	18	1.2	456	20	AAV86590	EST clone BG3. Ho
693	18	1.2	457	22	AA198810	Human excretory re
694	18	1.2	457	22	AA135597	Human musculoskele
695	18	1.2	457	22	AA163206	Human kidney relat
696	18	1.2	460	22	AAH34388	Human colon cancer
697	18	1.2	463	22	AA192884	Human polynucleoti
698	18	1.2	464	22	AA186425	Human polynucleoti
699	18	1.2	464	22	AA193001	Human polynucleoti
700	18	1.2	470	22	AAK66310	Human immune/haema
701	18	1.2	470	22	AAK66311	Human immune/haema
702	18	1.2	474	19	AAV59513	Human secreted pro
703	18	1.2	476	22	AA136728	Human musculoskele
704	18	1.2	477	22	ABA59173	Human foetal liver
705	18	1.2	477	22	ABA27943	Probe #6409 for ge
706	18	1.2	477	22	AAK07364	Human brain expres
707	18	1.2	477	22	AAK33131	Human bone marrow
708	18	1.2	477	22	AA116345	Probe #6278 for ge
709	18	1.2	477	22	AA138927	Probe #7613 used t
710	18	1.2	479	22	ABA42189	Human breast cell
711	18	1.2	479	22	ABA52611	Human foetal liver
712	18	1.2	479	22	ABA22399	Probe #865 for gen
713	18	1.2	479	22	AAK00872	Human brain expres
714	18	1.2	479	22	AAK26327	Human bone marrow
715	18	1.2	479	22	AA110960	Probe #893 for gen
716	18	1.2	479	22	AA132220	Probe #906 used to
717	18	1.2	479	22	AA100881	Probe #872 used to
718	18	1.2	481	22	AA116003	Human breast cancer
719	18	1.2	486	22	AA196933	Human Mammary Glan
720	18	1.2	489	22	AA113840	Human breast cancer
721	18	1.2	492	18	AA184952	Human prostate pro
722	18	1.2	492	20	AAK35882	CDNA encoding a pr
723	18	1.2	495	20	AA224821	Human secreted pro
724	18	1.2	496	22	AA111495	Human breast cancer
725	18	1.2	496	22	AA112047	Human breast cancer
726	18	1.2	499	21	AAK59372	Human secreted pro
727	18	1.2	501	22	AAK62927	Human immune/haema
728	18	1.2	503	22	AA113107	Human breast cancer
729	18	1.2	507	22	AAK35593	Human cardiovascular
730	18	1.2	508	22	ABA14314	Human nervous syst
731	18	1.2	508	22	ABA07689	Human ovarian and
732	18	1.2	508	22	AA102641	Human reproductive
733	18	1.2	515	22	AAH09428	Human CDNA clone (
734	18	1.2	515	22	ABA62628	Human foetal liver
735	18	1.2	516	22	AAK29942	Probe #8408 for ge
736	18	1.2	516	22	AAK10996	Human brain expres
737	18	1.2	516	22	AAK36838	Human bone marrow
738	18	1.2	516	22	AA117689	Probe #7622 for ge
739	18	1.2	516	22	AA142623	Probe #11309 used
740	18	1.2	517	22	AA180518	Human polynucleoti
741	18	1.2	519	20	AAK61880	Human immune/haema
742	18	1.2	521	20	AAV86963	EST clone B166. H
743	18	1.2	521	24	AA561743	Lung small cell ca
744	18	1.2	522	22	AA070907	Human secreted pro
745	18	1.2	523	22	AAK61793	Human immune/haema
746	18	1.2	526	16	AA086309	Wilson disease gen
747	18	1.2	527	22	AAK73020	Human immune/haema
748	18	1.2	529	22	AAK73018	Human immune/haema
749	18	1.2	530	22	AAK56473	Human CDNA for an
750	18	1.2	530	22	AAK73019	Human immune/haema
751	18	1.2	540	22	AA105347	Human reproductive
752	18	1.2	540	22	AAK71626	Human immune/haema
753	18	1.2	544	22	AAH10053	Human CDNA clone (
754	18	1.2	546	22	ABA60636	Human foetal liver
755	18	1.2	546	22	ABA28743	Probe #7209 for ge
756	18	1.2	546	22	AAK08918	Human brain expres
757	18	1.2	546	22	AAK34808	Human bone marrow
758	18	1.2	546	22	AA116835	Probe #6768 for ge
759	18	1.2	546	22	AA140527	Probe #9213 used t
760	18	1.2	548	22	AAH13288	Human CDNA clone (
761	18	1.2	548	22	AAK58546	Human immune/haema
762	18	1.2	552	22	AAK33132	Human secreted pro
763	18	1.2	555	22	AAH09476	Human CDNA clone (
764	18	1.2	555	22	AAH09942	Human CDNA clone (
765	18	1.2	560	22	AA102430	Human reproductive
766	18	1.2	560	22	AAK64677	Human immune/haema
767	18	1.2	564	22	AAH10298	Human CDNA clone (
768	18	1.2	567	22	AAK62114	Human immune/haema
769	18	1.2	567	22	AAK76252	Human immune/haema
770	18	1.2	567	22	AAK76253	Human immune/haema
771	18	1.2	568	22	AAK81769	Human immune/haema
772	18	1.2	569	22	AAH10316	Human CDNA clone (
773	18	1.2	570	22	AAH33656	Human colon cancer
774	18	1.2	571	21	AAK22330	Human secreted pro
775	18	1.2	573	22	AAK80470	Human immune/haema
776	18	1.2	575	22	AAK27483	CDNA encoding nove
777	18	1.2	582	21	AAH11948	Human cytokine rec
778	18	1.2	583	22	AAH10021	Human CDNA clone (
779	18	1.2	586	22	AAK71625	Human immune/haema
780	18	1.2	594	20	AAV88419	EST clone EK634.
781	18	1.2	597	22	AAK34743	Human DNA for a no
782	18	1.2	599	22	AAZ06774	Human phosphodiester
783	18	1.2	599	23	AA577758	DNA encoding novel
784	18	1.2	600	20	AAK30350	DNA encoding a hum
785	18	1.2	600	22	AA120932	Human breast cancer
786	18	1.2	604	22	AAH64831	Human secreted pro
787	18	1.2	604	22	AAK30544	DNA encoding novel
788	18	1.2	604	22	AA106320	Human reproductive
789	18	1.2	608	20	AAK20439	Human secreted pro
790	18	1.2	611	11	AAQ03848	Equine beta LH sub
791	18	1.2	612	22	AAK32346	Human CDNA encodin
792	18	1.2	618	22	AAK81393	Human immune/haema
793	18	1.2	619	21	AAK50674	Hermansky-Pudlak s
794	18	1.2	621	22	AAH34680	Human colon cancer
795	18	1.2	624	21	AAH16409	Human colon cancer
796	18	1.2	629	22	AAK75088	Human colon associ
797	18	1.2	633	22	ABA19495	Human nervous syst
798	18	1.2	634	22	AAK80474	Human immune/haema
799	18	1.2	635	22	AAK80472	Human immune/haema
800	18	1.2	636	21	AAK59368	Human secreted pro
801	18	1.2	641	23	AAK591349	DNA encoding novel
802	18	1.2	642	20	AAK73040	Human secreted pro
803	18	1.2	646	21	AAH81659	N. meningitidis pa
804	18	1.2	649	20	AAK24522	Human lung tumor a
805	18	1.2	649	21	AAK65761	Human lung cancer
806	18	1.2	650	22	AAK64518	Human immune/haema
807	18	1.2	652	22	ABA15611	Human nervous syst
808	18	1.2	652	23	AAK575704	DNA encoding novel
809	18	1.2	654	11	AAK04332	Plasmodium falcipar
810	18	1.2	657	22	AAK07935	Human TDC protein-
811	18	1.2	660	21	AAK96339	CDNA encoding a no
812	18	1.2	660	22	AAK56158	Human DNA encoding

813	18	1.2	664	23	AA572398	DNA encoding novel
814	18	1.2	665	22	AAH07112	Human cDNA clone (
815	18	1.2	666	22	AA684604	Human zrf4 protein
816	18	1.2	666	23	AA591699	DNA encoding novel
817	18	1.2	671	22	AA529154	CDNA encoding for
818	18	1.2	672	20	AA533959	Human HCMV repress
819	18	1.2	673	22	AA119259	Human breast cancer
820	18	1.2	675	22	AAK57397	Human immune/haema
821	18	1.2	675	19	AA652136	Leishmania donovan
822	18	1.2	676	18	AAV47556	Leishmania antigen
823	18	1.2	676	24	AA596020	Leishmania antigen
824	18	1.2	683	23	AAK64693	Human immune/haema
825	18	1.2	685	22	AA575708	DNA encoding novel
826	18	1.2	690	22	AAK87406	Human immune/haema
827	18	1.2	691	22	AA674447	Human PRO16 nucleio
828	18	1.2	692	22	AA102626	Human reproductive
829	18	1.2	692	22	AAAD08205	Human secreted pro
830	18	1.2	693	20	AA233643	Human breast tumou
831	18	1.2	694	19	AAV34166	Human secreted pro
832	18	1.2	694	19	AA526254	Human cDNA encodin
833	18	1.2	696	22	AA644906	Human breast cancer
834	18	1.2	699	22	ABAI7101	Human nervous syst
835	18	1.2	699	22	AAK59991	Human immune/haema
836	18	1.2	700	21	AA293153	Human CBCEP10 nuc
837	18	1.2	700	21	AA247519	CDNA sequence of w
838	18	1.2	700	22	AAH92268	Human inflammatory
839	18	1.2	700	22	AAH92269	Human inflammatory
840	18	1.2	700	22	AAH92617	Human inflammatory
841	18	1.2	700	22	AAH92618	Human inflammatory
842	18	1.2	700	22	AAH93072	Human inflammatory
843	18	1.2	706	21	AAAC74313	Human secreted pro
844	18	1.2	706	22	AAH64918	Human secreted pro
845	18	1.2	708	22	AA120885	Human breast cancer
846	18	1.2	709	22	AAK85680	Human immune/haema
847	18	1.2	711	22	AAK85689	Human immune/haema
848	18	1.2	720	22	ABAO7921	Human ovarian and
849	18	1.2	720	22	AA103738	Human reproductive
850	18	1.2	720	22	AAH04566	Human immune/haema
851	18	1.2	720	22	AAH07620	Human cDNA clone (
852	18	1.2	721	20	AAH88004	Human secreted pro
853	18	1.2	724	23	AA591434	DNA encoding novel
854	18	1.2	725	20	AAK37532	Human secreted pro
855	18	1.2	725	21	AA195619	Human neuroblastom
856	18	1.2	726	21	AAAC76265	Human ORFX ORF1820
857	18	1.2	726	22	AA533105	DNA encoding human
858	18	1.2	727	20	AAK20487	Human secreted pro
859	18	1.2	727	22	AA527247	CDNA encoding nove
860	18	1.2	729	21	AAFL3585	Aspergillus oryzae
861	18	1.2	736	23	AA586863	DNA encoding novel
862	18	1.2	737	22	AA121977	Human breast cancer
863	18	1.2	744	22	AA195742	Human neuroblastom
864	18	1.2	749	22	AAH64931	Human secreted pro
865	18	1.2	750	20	AAZ16810	Human gene express
866	18	1.2	750	20	AAZ16487	Human gene express
867	18	1.2	751	22	AAK76824	Human immune/haema
868	18	1.2	753	20	AAZ00347	Nucleotide sequenc
869	18	1.2	753	22	AA689005	Human FATP3 coding
870	18	1.2	755	20	AAZ16214	Human gene express
871	18	1.2	757	20	AAZ16857	Human gene express
872	18	1.2	757	20	AAZ16858	Human gene express
873	18	1.2	757	22	AAH04118	Human cDNA clone (
874	18	1.2	758	22	AA526521	Human cDNA encodin
875	18	1.2	761	22	AAH70851	Human cervical can
876	18	1.2	763	22	AA686932	Nucleotide sequenc
877	18	1.2	769	22	AA113748	Human breast cancer
878	18	1.2	775	21	AAAC61101	Human cDNA encodin
879	18	1.2	775	22	AA195644	Human neuroblastom
880	18	1.2	776	20	AAZ52955	Human prostate tum
881	18	1.2	777	21	AAAC99821	Human secreted pro
882	18	1.2	778	21	AAZ24830	Human secreted pro
883	18	1.2	781	22	AA124847	Human breast cancer
884	18	1.2	783	21	AAAC74316	Human secreted pro
885	18	1.2	783	22	AAK77935	Human immune/haema
886	18	1.2	787	13	AAO27946	Sequence encoding
887	18	1.2	791	21	AAAC74347	Human secreted pro
888	18	1.2	794	21	AAK89446	Corn branched chal
889	18	1.2	794	22	AA122617	Human breast cancer
890	18	1.2	794	22	AAK91740	Human cDNA 5'-end
891	18	1.2	794	22	AAK93760	Human cDNA clone r
892	18	1.2	797	20	AAK85547	Canine hookworm ne
893	18	1.2	797	22	AAH13386	Human secreted pro
894	18	1.2	804	22	AAH64840	Human secreted pro
895	18	1.2	806	19	AAV04568	Filea serine protea
896	18	1.2	806	22	AAAC90875	Arabidopsis thalia
897	18	1.2	809	21	AAAC51789	Partial human SLC6
898	18	1.2	812	22	AA673814	Human reproductive
899	18	1.2	813	22	AA104356	Human reproductive
900	18	1.2	813	22	AA104359	Human breast cancer
901	18	1.2	814	22	AA118571	Human secreted pro
902	18	1.2	829	22	AA502450	Human secreted pro
903	18	1.2	830	22	AAK79552	Human immune/haema
904	18	1.2	831	20	AAV84475	Human secreted pro
905	18	1.2	831	22	ABAB3258	Human secreted pro
906	18	1.2	835	21	AAAC53455	Arabidopsis thalia
907	18	1.2	836	21	AA531289	Human cDNA encodin
908	18	1.2	839	21	AAAC59590	Human secreted pro
909	18	1.2	839	22	AAK82054	Human immune/haema
910	18	1.2	845	21	AAAC40933	Arabidopsis thalia
911	18	1.2	849	20	AAK27372	Human secreted pro
912	18	1.2	854	21	AAAO2630	Human colon cancer
913	18	1.2	859	18	AA162944	Human eosinocyte C
914	18	1.2	872	22	AA674195	DNA encoding eviro
915	18	1.2	877	11	AAO03999	Sequence complemen
916	18	1.2	881	20	AAZ16635	Human gene express
917	18	1.2	886	22	AA197734	Human neuroblastom
918	18	1.2	887	21	AAA26435	Human secreted pro
919	18	1.2	893	21	AA681794	Human secreted pro
920	18	1.2	900	21	AAAC55216	Human secreted pro
921	18	1.2	904	22	AAK85179	Human immune/haema
922	18	1.2	904	22	AAK85180	Human immune/haema
923	18	1.2	906	22	AA194153	Human neuroblastom
924	18	1.2	906	22	AA502228	Soybean Wuschel (W
925	18	1.2	913	20	AAK98003	Human secreted pro
926	18	1.2	920	21	AAAC99934	Human secreted pro
927	18	1.2	925	20	AAK920440	Human secreted pro
928	18	1.2	925	22	AAH94428	Human foetal cDNA,
929	18	1.2	931	21	AAAC59123	Human secreted pro
930	18	1.2	936	22	AA197770	Human neuroblastom
931	18	1.2	939	21	AAAC59753	Human secreted pro
932	18	1.2	943	20	AAZ00438	Human secreted pro
933	18	1.2	945	22	AAH72738	Human secreted pro
934	18	1.2	947	22	AAH77061	Human cervical can
935	18	1.2	948	18	AAV74991	Proliferative glom
936	18	1.2	948	18	AAAD08462	Staphylococcus aur
937	18	1.2	954	18	AA161031	Human secreted pro
938	18	1.2	960	21	AAV6279	Aureobasidium pull
939	18	1.2	966	21	AAAL05751	Maize glutathione-
940	18	1.2	967	22	AAK74532	Human reproductive
941	18	1.2	970	21	AAAC36516	Human immune/haema
942	18	1.2	982	20	AAK37417	Arabidopsis thalia
943	18	1.2	982	20	AAK68151	Human secreted pro
944	18	1.2	1000	20	AAZ27734	Human immune/haema
945	18	1.2	1000	21	AAAA47308	Human DNA marker c
946	18	1.2	1003	21	AAAC79708	Human genomic DNA
947	18	1.2	1006	22	AAH64824	Human secreted pro
948	18	1.2	1006	22	AAH64824	Human secreted pro
949	18	1.2	1011	15	AAO58038	Interferon gamma g
950	18	1.2	1015	18	AA772166	Alzheimer's disease
951	18	1.2	1016	22	AAH32563	Human secreted pro
952	18	1.2	1031	22	AAAC76779	Human ORFX ORF2334
953	18	1.2	1031	22	AA536431	Human cardiovascular
954	18	1.2	1035	21	AA534169	Human cDNA encodin
955	18	1.2	1036	21	AAK93920	C-terminal portion
956	18	1.2	1039	20	AAZ00423	Human secreted pro
957	18	1.2	1049	22	AAAC80651	Human secreted pro
958	18	1.2	1050	23	AAH31354	Human secreted pro
					AA575691	DNA encoding novel

959	18	1.2	1061	21	AAA26423	Human secreted pro
960	18	1.2	1062	21	AAC69475	Human secreted pro
961	18	1.2	1066	22	AAD04001	Human protein tyro
962	18	1.2	1066	24	AAS62527	CDNA sequence #314
963	18	1.2	1070	24	AAS62527	CDNA sequence #436
964	18	1.2	1084	21	AAC59873	Human secreted pro
965	18	1.2	1089	21	AAC99677	Human secreted pro
966	18	1.2	1089	23	AAS87173	DNA encoding novel
967	18	1.2	1094	21	AAC59923	Human secreted pro
968	18	1.2	1098	21	AAC79709	Human secreted pro
969	18	1.2	1100	17	AAT40216	Sequence encoding
970	18	1.2	1101	21	AAC77676	Human cancer assoc
971	18	1.2	1102	19	AAV28734	Murine smooth musc
972	18	1.2	1102	22	AAD21105	Murine SM22 alpha
973	18	1.2	1102	22	AAD20564	Murine smooth musc
974	18	1.2	1102	22	AAB26685	Mouse smooth muscl
975	18	1.2	1105	22	AAB17630	Human nervous syst
976	18	1.2	1107	21	AAC77646	Human cancer assoc
977	18	1.2	1109	22	AAS40964	CDNA encoding nove
978	18	1.2	1111	21	AAC74346	Human secreted pro
979	18	1.2	1113	21	AAZ97105	Human secreted pro
980	18	1.2	1114	24	AAS62383	CDNA sequence #170
981	18	1.2	1121	22	AAD07933	Human TDC protein-
982	18	1.2	1125	22	ABA06543	Human CDNA SEQ ID
983	18	1.2	1130	22	AAZ24818	Human secreted pro
984	18	1.2	1130	22	AAI06810	Human reproductive
985	18	1.2	1130	22	AAI62691	Human breast or ov
986	18	1.2	1132	22	AAI06808	Human reproductive
987	18	1.2	1132	22	AAI62689	Human breast or ov
988	18	1.2	1137	22	AAK71677	Human immune/haema
989	18	1.2	1149	22	AAK32117	Human secreted pro
990	18	1.2	1160	21	AAC39976	Arabidopsis thalia
991	18	1.2	1161	24	ABL01566	Human secreted pro
992	18	1.2	1165	18	AAT72945	Phaffia carotenoid
993	18	1.2	1167	22	AAF64199	Human secreted pro
994	18	1.2	1181	24	AAS62665	CDNA sequence #452
995	18	1.2	1189	20	AAZ00813	Human secreted pro
996	18	1.2	1201	21	AAC44975	Arabidopsis thalia
997	18	1.2	1205	21	AAC33121	Arabidopsis thalia
998	18	1.2	1210	24	ABA92619	Human class II ami
999	18	1.2	1228	21	AAC80567	Human secreted pro
1000	18	1.2	1238	22	AAK65390	Human immune/haema

## ALIGNMENTS

## RESULT 1

AAH20992 standard; cDNA; 1539 BP.

AAH20992;  
31-AUG-2001 (first entry)  
D. melanogaster acetylcholine receptor beta subunit Db3 cDNA.  
Acetylcholine receptor beta subunit; Db3; fruitfly; insecticidal;  
transgenic; modulator; plant-protection agent; medicine; ss.  
Drosophila melanogaster.  
Key Location/Qualifiers  
CDS 43..1368  
/tag= a  
/product= "Db3"  
EP1106689-A2.  
13-JUN-2001.  
28-NOV-2000; 2000EP-0125300.

PR 10-DEC-1999; 99DE-1059582.  
XX (FARB ) BAYER AG.  
PA Adamczewski M, Methfessel C, Schulte T;  
XX WPI; 2001-376906/40.  
DR P-PDB; AAB86218.  
XX New nucleic acid encoding Drosophila acetylcholine receptor subunit,  
PT useful in screening for plant protection agents and pharmaceuticals.  
PS Claim 1; Page 10-13; 18pp; German.  
XX  
XX This invention describes a novel nucleic acid (I) which encodes a  
CC Drosophila melanogaster acetylcholine receptor beta subunit, Db3 which  
CC has insecticidal activity. (II), related host cells, polypeptides, ACR,  
CC antibodies, transgenic invertebrates and regulatory regions, are used to  
CC identify agents that modulate ACR activity, potentially useful as  
CC plant-protection agents (e.g. insecticides) or pharmaceuticals for human  
CC or veterinary medicine. This sequence encodes the Drosophila melanogaster  
CC acetylcholine receptor beta subunit described in the method of the  
CC invention.  
XX  
XX Sequence 1539 BP; 352 A; 476 C; 411 G; 300 T; 0 other;  
SQ

Query Match 100.0%; Score 1539; DB 22; Length 1539;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atccgcacgaggtatcatccgaacaaagcgcgctgaacaatgacgactcccaag 60  
DB 1 atccgcacgaggtatcatccgaacaaagcgcgctgaacaatgacgactcccaag 60  
QY 61 ataaagcaccagttccggtcctgactgacactactgtgtaaatgctaaaggagat 120  
DB 61 ataaagcaccagttccggtcctgactgacactactgtgtaaatgctaaaggagat 120  
QY 121 ctcttatggggtgacttcggtgacgagcgacatgcccacggaccgaagcgc 180  
DB 121 ctcttatggggtgacttcggtgacgagcgacatgcccacggaccgaagcgc 180  
QY 121 ctcttatggggtgacttcggtgacgagcgacatgcccacggaccgaagcgc 180  
DB 121 ctcttatggggtgacttcggtgacgagcgacatgcccacggaccgaagcgc 180  
QY 181 aatgtcaaggccctgtgatcgcctcaccgcccgtgttcaagaactacgacgagtgt 240  
DB 181 aatgtcaaggccctgtgatcgcctcaccgcccgtgttcaagaactacgacgagtgt 240  
QY 241 caagcggtgttccaaggaaccccaacgaacggtgctcctggaatgtgtacactacata 300  
DB 241 caagcggtgttccaaggaaccccaacgaacggtgctcctggaatgtgtacactacata 300  
QY 301 gacatgcagcaggttgaacgcaagctgacacccactgtcgtgatactccgatggaga 360  
DB 301 gacatgcagcaggttgaacgcaagctgacacccactgtcgtgatactccgatggaga 360  
QY 361 gacgaggaagcggtgtgcaacccgtcaacaatgacaaacacacgacgacacttgaag 420  
DB 361 gacgaggaagcggtgtgcaacccgtcaacaatgacaaacacacgacgacacttgaag 420  
QY 421 tccagcgaggtctggaaccccaatcacacttcaacgagcagcagagtgctcagcgatg 480  
DB 421 tccagcgaggtctggaaccccaatcacacttcaacgagcagcagagtgctcagcgatg 480  
QY 481 gccgaacccaggtgacccctcagcagacgacatgacacttcggtgagtgctcagcgatg 540  
DB 481 gccgaacccaggtgacccctcagcagacgacatgacacttcggtgagtgctcagcgatg 540  
QY 541 tacacggtactatgcgaactcaacatgctcaactggtcccaagcagacgagctgcaag 600  
DB 541 tacacggtactatgcgaactcaacatgctcaactggtcccaagcagacgagctgcaag 600  
QY 601 ttgaagatcggtcctcggtggtcgtgaaggtgctcctcggagagaagcagcgagagaaga 660  
DB 601 ttgaagatcggtcctcggtggtcgtgaaggtgctcctcggagagaagcagcgagagaaga 660



Db 601 ttgaagatcggtctcttggggtctgaaggtctcttgcggagaaagcagcgagagaga 660  
 Qy 661 gattcccttgacacagcagcgttgcagtcacccgagatgggaatcgttgaactcgcga 720  
 Db 661 gattcccttgacacagcagcgttgcagtcacccgagatgggaatcgttgaactcgcga 720  
 Qy 721 gcccaacttgcagtcagcagcactacacgctacatgagatgacactcgcagcgcga 780  
 Db 721 gcccaacttgcagtcagcagcactacacgctacatgagatgacactcgcagcgcga 780  
 Qy 781 cgtctccatgtacacagcgcgtacatcacacccgcgtctcgtcatctcgtcc 840  
 Db 781 cgtctccatgtacacagcgcgtacatcacacccgcgtctcgtcatctcgtcc 840  
 Qy 841 ctccacagcttctggtcgtcccccacatgagcgagagagatcatgataagcgtc 900  
 Db 841 ctccacagcttctggtcgtcccccacatgagcgagagagatcatgataagcgtc 900  
 Qy 901 ctcatcatcgtacacgcgcctctcctcatgtactcgcagcctcctcgtcagtgctcc 960  
 Db 901 ctcatcatcgtacacgcgcctctcctcatgtactcgcagcctcctcgtcagtgctcc 960  
 Qy 1081 gaagcgtctgagaaagcgtctgacagcgacacgtgagcagcgtcgtcgtcttc 1140  
 Db 1081 gaagcgtctgagaaagcgtctgacagcgacacgtgagcagcgtcgtcgtcttc 1140  
 Qy 1141 agcaccaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1200  
 Db 1141 agcaccaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1200  
 Qy 1201 gaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1260  
 Db 1201 gaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1260  
 Qy 1261 aagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1320  
 Db 1261 aagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1320  
 Qy 1321 tctctcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1380  
 Db 1321 tctctcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1380  
 Qy 1381 ctcaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1440  
 Db 1381 ctcaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1440  
 Qy 1441 gttattatgt 1500  
 Db 1441 gttattatgt 1500  
 Qy 1501 gacctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1539  
 Db 1501 gacctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1539

RESULT 2

ID ABL23317 standard; DNA: 1362 BP.

AC ABL23317;

DF 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 21424.

KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ds.  
 OS Drosophila melanogaster.  
 PN WO200171042-A2.  
 PD 27-SEP-2001.  
 PF 23-MAR-2001; 2001WO-US09231.  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 PA (PEKE ) PE CORP NY.  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 DR WPI; 2001-656860/75.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 PS  
 PS Claim 1: SEQ ID NO 21424; 21pp + Sequence Listing; English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins  
 CC (AB101840-AB116175).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX  
 XX  
 SQ Sequence 1362 BP; 294 A; 444 C; 374 G; 250 T; 0 other;

Query Match 78.1%; Score 1202; DB 23; Length 1362;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1352; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 14 gtacatccgaacaaagcgcgtcgaacatgacagcgtcctccaagataaaggcaccag 73  
 Db 8 gtacatccgaacaaagcgcgtcgaacatgacagcgtcctccaagataaaggcaccag 67  
 Qy 74 tttcgggtcgtgagctgccaactatgtctgcaaatgcaatgaggtatctctttatggggc 133  
 Db 68 tttcgggtcgtgagctgccaactatgtctgcaaatgcaatgaggtatctctttatggggc 127  
 Qy 134 tgacttcgtgtcagcgcgcacttgcacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 193  
 Db 128 tgacttcgtgtcagcgcgcacttgcacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 187  
 Qy 194 tggatcgtctcgaacgcgcgtgttcaagcactcgaacgcgcgcgcgcgcgcgcgcgcgcgc 253  
 Db 188 tggatcgtctcgaacgcgcgtgttcaagcactcgaacgcgcgcgcgcgcgcgcgcgcgcgc 247  
 Qy 254 aaggaaccccccagaacgtgtcccttggaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 313  
 Db 248 aaggaaccccccagaacgtgtcccttggaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 307  
 Qy 314 tgaacgcaagctgac 373  
 Db 308 tgaacgcaagctgac 367  
 Qy 374 tgtgcaacgttcaatatagacaacatcagcagatcactttgaatcagcaggtct 433  
 Db 368 tgtgcaacgttcaatatagacaacatcagcagatcactttgaatcagcaggtct 427





PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Claim 1: SEQ ID NO 21409; 21bp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA  
CC sequences (ABU161840-ABU16175) and the encoded proteins  
CC (ABU57737-ABU72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 3764 BP; 943 A; 972 C; 960 G; 889 T; 0 other;

Query Match 12.7%; Score 195; DB 23; Length 3764;  
Best Local Similarity 100.0%; Pred No. 4.5e-72;  
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 gccaccgcgcagccccaagaagcccaatgtcaagccctgacccccaagccgcgcctg 216  
DB 271 GCCACGCGGAGCCCAAGAGCGCAATGTCAAGCCCTGATCGCTCCACGCCGCGCTG 212  
QY 217 ttttgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 276  
DB 211 TTACGCAACTACGACAGCGATGTGACGCCGTCTCCAAAGAACCCCGACGATGTCC 152  
QY 277 ctggaatggtgtcaccctacatagacatcagcagcagcagcagcagcagcagcagcagcagc 336  
DB 151 CTGGAATGTGTCTTACCTATACATAGACATGAGAGATGTAACGCAAGCTGACACCCAC 92  
QY 337 tgcgtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 351  
DB 91 TGCTGCTGAATCTC 77

RESULT 6  
AAL36376  
ID AAL36376 standard; DNA; 571 BP.  
XX  
AC AAL36376;  
XX  
DE 08-JAN-2002 (first entry)  
XX  
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2741.  
XX  
KM Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KM anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;  
KM vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KM neurological disease; infection; human; secreted protein;  
KM musculoskeletal system; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155367-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001MO-US01338.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0231415.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.

XX 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244517.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249267.  
PR 17-NOV-2000; 2000US-0249289.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251088.  
PR 05-DEC-2000; 2000US-0251938.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251889.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX WPI: 2001-451937/48.  
XX  
XX Isolated polypeptide for treating, preventing and/ or prognosing  
PT disorders related to the musculoskeletal system including  
PT musculoskeletal cancers and also for testing and detection e.g.  
PT diagnosis -  
XX  
XX Example 2; SEQ ID NO 2741; 781pp + Sequence Listing; English.  
XX  
CC The invention relates to novel genes (AAL34669-AAL37666) and proteins  
(AAB03087-AAB04109) associated with the musculoskeletal system useful

CC for preventing, treating or ameliorating medical conditions e.g. by  
CC protein or gene therapy. The genes are isolated from a range of human  
CC tissues disclosed in the specification. The nucleic acids, proteins,  
CC antibodies and (ant)agonists are useful in the diagnosis, treatment  
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and  
CC other cancers of the adrenal gland, bone, bone marrow, breast,  
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound  
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;  
CC and (f) infectious diseases such as viral, bacterial, fungal and  
CC parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at [ftp.wipo.int/pub/published\\_pat\\_sequences](http://ftp.wipo.int/pub/published_pat_sequences).  
CC  
CC  
CC Sequence 571 BP; 157 A; 137 C; 118 G; 159 T; 0 other;

Query Match	1.6%;	Score 24;	DB 22;	Length 571;
Best Local Similarity	100.0%;	Pred. No. 2;		
Matches 24;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	1516	tctgcttacaataaaaaaaa	1539	
db	490	tctgcttacaataaaaaaaa	513	

RESULT	7
AAH41137	
ID	AAH41137 standard; DNA; 1135 BP.

DT 21-AUG-2001 (first entry)

DE Arabidopsis thaliana chlorophyllase coding sequence #2.

KW Chlorophyllase; transgenic plant; ds.

OS *Arabidopsis thaliana*.

JP2001086990-A.

PD 03-APR-2001.

PF 20-SEP-1999; 99JP-0266181.

PR 20-SEP-1999; 99JP-0266181.

PA (KAGO ) KAGOME KK.

AA  
DR WPI: 2001-338421/36.

DR E-PSDB; AAB99103.  
XX

PI DNA encoding chlorophyllase, useful for producing transgenic plants - XX

PS Claim 5; page 12-14; 21pp; Japanese.

CC The present sequence is a chlorophyllase protein coding sequence. The

entrophyase protein coding sequence can be used for the transformation of a plant.

Sequence 1135 BP; 331 A; 233 C; 247 G; 324 T; 0 other;

Query Match	1.4%	Score 22;	DB 22;	length 1135;
Best Local Similarity	100.0%	Pred. No. 13;		
Matches	22;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps	0;			
OY	1518	tgcttaacaaaaaaaaa	1539	

Db 1106 Tgcttacaataaaaaaaa 1127

RESULT 8

AAZ24886  
ID AAZ24886 standard; DNA; 1377 BP.  
XX  
XX AAZ24886;  
AC  
XX  
XX 02-DEC-1999 (first entry)  
DT  
XX  
XX Human secreted protein gene 76 clone HFPBM30.  
DE

Human; secreted protein; fusion protein; gene therapy; protein therapy;  
diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

XX W09947540-A1.

XX 23-SEP-1999.

XX 18-MAR-1999; 99WO-US05804.

XX 19-MAR-1998; 98US-0078563.  
XX 19-MAR-1998; 98US-0078566.  
XX 19-MAR-1998; 98US-0078573.  
XX 19-MAR-1998; 98US-0078574.  
XX 19-MAR-1998; 98US-0078576.  
XX 19-MAR-1998; 98US-0078577.  
XX 19-MAR-1998; 98US-0078578.  
XX 19-MAR-1998; 98US-0078579.  
XX 19-MAR-1998; 98US-0078581.  
XX 01-APR-1998; 98US-0080312.  
XX 01-APR-1998; 98US-0080313.  
XX 01-APR-1998; 98US-0080314.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;  
PI Wei Y, Endress GA, Duan RD, Kyaw H, Ebner R, Lafleur DM;  
PI Olsen HS, Shi Y, Moore PA;

DR WPI; 1999-562050/47.  
DR P-PSDB; AAY41383.

PT New isolated human genes, useful for diagnosis and treatment of e.g.  
PT cancers, neurological disorders, immune diseases, inflammation or blood  
PT disorders

PS Claim 1; Page 343-344; 484pp; English.

CC This sequence represents a nucleic acid molecule which encodes a  
CC secreted human protein. The gene number, and the clone it is derived  
CC from, are detailed in the descriptor line. The gene can be used to  
CC generate fusion proteins by linking to the gene to a human immunoglobulin  
CC Fc portion (e.g. AAZ24802) for increasing the stability of the fused  
CC protein as compared to the human protein only.

CC The invention relates to 95 novel genes and their fragments (nucleic  
CC acid sequences: AAZ24811-224907; amino acid sequences AAY41308-741404)  
CC which are useful for preventing, treating or ameliorating medical  
CC conditions e.g. by protein or gene therapy. Also, pathological  
CC conditions can be diagnosed by determining the amount of the new  
CC polypeptides in a sample or by determining the presence of mutations in  
CC the new polynucleotides. Specific uses are described for each of the 95  
CC polynucleotides, based on which tissues they are most highly expressed in

CC (see AAZ24811 for described uses).

XX Sequence 1377 BP; 311 A; 283 C; 379 G; 404 T; 0 other;

Query Match 1.4%; Score 21; DB 20; Length 1377;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1519 gcttacaataaaaaaaa 1539  
DB 1354 gcttacaataaaaaaaa 1374

RESULT 9

AAH32948  
ID AAH32948 standard; cDNA; 1843 BP.

XX AAH32948;  
AC

XX 03-SEP-2001 (first entry)  
DT

XX Human colon cancer antigen encoding cDNA SEQ ID NO:4.  
DE

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
XX colorectal carcinoma; ss.

OS Homo sapiens.

XX W0200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 98US-0157137.  
XX 03-NOV-1999; 99US-0163280.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI; 2001-235357/24.  
DR P-PSDB; AAG73517.

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers -

PS Claim 1; Page 2231-2232; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytosolic activity and can be used in gene  
CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing  
CC inactive proteins or to supplement the patient's own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated P,  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC to express the proteins. N and P can be used in the prevention, diagnosis  
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
CC and AAG77789 represent sequences used in the exemplification of the  
CC present invention.

CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 1843 BP; 632 A; 329 C; 310 G; 562 T; 10 other;



DT 24-JAN-2002 (first entry)  
XX  
DE Human SNP oligonucleotide #4635.  
XX  
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;  
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KW multifactorial disease; autoimmune disease; infection;  
KW nervous system disease; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200147944-A2.  
XX  
PD 05-JUL-2001.  
XX  
PF 28-DEC-2000; 2000WO-US35498.  
XX  
PR 28-DEC-1999; 99US-0173419.  
XX  
PR 27-DEC-2000; 2000US-0173419.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Shinkets RA, Leach M;  
XX  
PI WPI; 2001-465210/50.  
XX  
DR Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
XX  
PT oncogenes and histones, useful for diagnosing and treating, e.g.  
XX  
PT cancer, autoimmune diseases and infections -  
XX  
PS Claim 1; Page 2719; 4143pp; English.  
XX  
CC The present invention relates to oligonucleotides encoding polymorphic  
CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
CC histones, kinases, colony stimulating factors, complement related  
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,  
CC G-protein coupled receptors and thioesterases. The present sequence is  
CC one such oligonucleotide. The oligonucleotides and the peptides encoded  
CC by them may be used in the prevention, diagnosis and treatment of  
CC diseases associated with inappropriate expression of the proteins listed  
CC above. Disorders that may be prevented, diagnosed and/or treated include  
CC multifactorial diseases with a genetic component, such as autoimmune  
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
CC leukaemia), diseases of the nervous system and an infection of pathogenic  
CC organisms.  
XX  
CC  
SQ Sequence 51 BP; 26 A; 7 C; 8 G; 10 T; 0 other;  
XX  
XX  
Query Match 1.3%; Score 20; DB 22; Length 51;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1519 gcttacaataaaaaaaa 1538  
DB 20 gcttacaataaaaaaaa 39

RESULT 13  
ALU31866/C  
ID AAL31866 standard; DNA; 51 BP.  
XX  
AC AAL31866;  
XX  
DT 24-JAN-2002 (first entry)  
XX

DE Human SNP oligonucleotide #5074.  
XX  
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;  
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KW multifactorial disease; autoimmune disease; infection;  
KW nervous system disease; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200147944-A2.  
XX  
PD 05-JUL-2001.  
XX  
PF 28-DEC-2000; 2000WO-US35498.  
XX  
PR 28-DEC-1999; 99US-0173419.  
XX  
PR 27-DEC-2000; 2000US-0173419.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Shinkets RA, Leach M;  
XX  
PI WPI; 2001-465210/50.  
XX  
DR Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
XX  
PT oncogenes and histones, useful for diagnosing and treating, e.g.  
XX  
PT cancer, autoimmune diseases and infections -  
XX  
PS Claim 1; Page 2846; 4143pp; English.  
XX  
CC The present invention relates to oligonucleotides encoding polymorphic  
CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
CC histones, kinases, colony stimulating factors, complement related  
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,  
CC G-protein coupled receptors and thioesterases. The present sequence is  
CC one such oligonucleotide. The oligonucleotides and the peptides encoded  
CC by them may be used in the prevention, diagnosis and treatment of  
CC diseases associated with inappropriate expression of the proteins listed  
CC above. Disorders that may be prevented, diagnosed and/or treated include  
CC multifactorial diseases with a genetic component, such as autoimmune  
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
CC leukaemia), diseases of the nervous system and an infection of pathogenic  
CC organisms.  
XX  
CC  
SQ Sequence 51 BP; 9 A; 5 C; 5 G; 32 T; 0 other;  
XX  
XX  
Query Match 1.3%; Score 20; DB 22; Length 51;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1520 cttaacataaaaaaaa 1539  
DB 30 cttaacataaaaaaaa 11

RESULT 14  
AAS56571  
ID AAS56571 standard; cDNA; 217 BP.  
XX  
AC AAS56571;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Human cDNA for an ovarian cancer protein #195.  
XX



KW Human; ss; ovarian cancer protein; cancer; tumour; ovarian cancer;  
KW endometrial cancer; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO200170976-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 20-MAR-2001; 2001WO-US09062.  
XX  
XX 21-MAR-2000; 2000US-190710P.  
PR 22-JUN-2000; 2000US-213748P.  
PR 19-DEC-2000; 2000US-257276P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
PI Xu J, Pyle RA, Stolk JA;  
XX  
XX WPI; 2001-607531/69.  
XX  
XX Nucleic acids encoding 222 polypeptides associated with ovarian and  
PT endometrial cancers, useful for diagnosing, preventing and treating  
PT cancers -  
XX  
PS Claim 1; Page 177; 187pp; English.  
XX  
XX The invention relates to human polynucleotides encoding proteins  
CC associated with ovarian and endometrial cancers. The polynucleotides and  
CC the proteins they encode may be used in the prevention, diagnosis and  
CC treatment of diseases associated with the inappropriate expression of  
CC ovarian and endometrial cancer polypeptides (OECs). For example, the  
CC polynucleotide (or an expression vector comprising the polynucleotide)  
CC and the OECs may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of OECs by expressing inactive proteins or to  
CC supplement the patients own production of them. Additionally, the  
CC polynucleotide may be used to produce the OECs, by inserting the nucleic  
CC acids into a host cell and culturing the cell to express the protein. The  
CC polynucleotide and its complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar nucleic acids in samples, and therefore which patients may be in  
CC need of restorative therapy. The OECs may also be used as antigens in  
CC the production of anti-OEC antibodies and in assays to identify  
CC modulators of its expression and activity. The anti-OEC antibodies and  
CC antagonists may also be used to down regulate expression and activity.  
CC The anti-OEC antibodies may also be used as diagnostic agents for  
CC detecting the presence of OECs in samples (e.g. by enzyme linked  
CC immunosorbant assay (ELISA)) and hence diagnose patients with  
CC cancers. The present sequence is a ovarian and endometrial cancer linked  
CC cDNA of the invention.  
XX  
SQ Sequence 217 BP; 85 A; 36 C; 37 G; 59 T; 0 other;

Query Match 1.3%; Score 20; DB 22; Length 217;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1520 cttacaaaaaaaaaaaaa 1539  
|||||  
DB 197 cttacaaaaaaaaaaaaa 216

RESULT 15  
AAS24909  
ID AAS24909 standard; CDNA: 290 BP.  
XX  
AC AAS24909;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human ovarian PCR-subtracted cDNA library clone #1090.

XX  
KW Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;  
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;  
KW primer; probe.  
XX  
OS Homo sapiens.  
XX  
PN WO200157207-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 05-FEB-2001; 2001WO-US03733.  
XX  
XX 04-FEB-2000; 2000US-0180403.  
PR 28-MAR-2000; 2000US-0192745.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
PI Algate PA, Mannion J;  
XX  
XX WPI; 2001-48879/53.  
XX  
XX New polynucleotides encoding ovarian tumour proteins, useful for  
PT treating ovarian cancer, and as probes, primers, and markers of cancer  
PT progression -  
XX  
PS Example 1; page 282; 378pp; English.  
XX  
XX The invention comprises compositions used for the therapy and diagnosis  
CC of ovarian cancer. The compositions comprise one or more ovarian tumour  
CC proteins, their associated polynucleotides, or immunogenic portions of  
CC the proteins. The ovarian tumour polynucleotides and polypeptides are  
CC useful for stimulating and/or expanding T cells specific for a tumour  
CC protein. They are also useful for inhibiting the development of cancer in  
CC a patient with an ovarian tumour DNA or protein by incubating isolated  
CC T-cells allowing them to proliferate, and administering to the patient.  
CC The sequences can be used as markers for cancer, for example, to monitor  
CC ovarian cancer progression. Probes and primers are useful in nucleic acid  
CC hybridisation, in detecting the presence of complementary sequences in a  
CC given sample, for preparing mutant species and for preparing other  
CC genetic constructions. Sequences AAS23820-AAS25231 and AAS25328-AAS25549  
CC represent human ovarian tumour protein cDNA clones.  
XX  
SQ Sequence 290 BP; 99 A; 90 C; 35 G; 66 T; 0 other;

Query Match 1.3%; Score 20; DB 22; Length 290;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1520 cttacaaaaaaaaaaaaa 1539  
|||||  
DB 269 cttacaaaaaaaaaaaaa 288

Search completed: July 18, 2002, 11:54:04  
Job time: 5227 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 10:18:12 ; Search time 132.17 Seconds  
(without alignments)  
577.217 Million cell updates/sec

Title: US-09-732-680A-2  
Perfect score: 2311  
Sequence: 1 MTTTPKIKAPVSGPLLL.....RISFVSFLAFLIIRCSV 441

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP\_TREMBL\_19:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_plant:\*  
11: sp\_protent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	2311	100.0	441 5 Q9VP08	Q9VP08 drosophila
2	1650	71.4	327 5 Q9SR44	Q9SR44 drosophila
3	434	18.8	512 11 Q91X60	Q91X60 mus musculu
4	418.5	18.1	457 6 Q9XS62	Q9XS62 canis famli
5	417	18.0	452 13 Q91RC7	Q91RC7 gallus gall
6	417	18.0	629 11 Q9EM51	Q9EM51 mus musculu
7	412	17.8	629 11 Q923N8	Q923N8 mus musculu
8	412	17.8	641 11 Q923N7	Q923N7 mus musculu
9	411.5	17.8	664 5 Q9N508	Q9N508 caenorhabdi
10	406	17.6	484 4 Q9BS29	Q9BS29 homo sapien
11	404.5	17.5	447 11 Q9JLB5	Q9JLB5 rattus norv
12	398	17.2	494 11 Q9ROM9	Q9ROM9 mus musculu
13	397.5	17.2	502 5 Q96AN6	Q96AN6 caenorhabdi
14	395	17.1	523 5 Q46128	Q46128 heliothis v
15	392.5	17.0	501 4 Q96F88	Q96F88 homo sapien
16	389	16.8	461 5 P91197	P91197 caenorhabdi

17	388.5	16.8	537 5 Q9U941	Q9U941 myzus persi
18	386	16.7	795 5 Q18394	Q18394 drosophila
19	385.5	16.7	545 5 Q96631	Q96631 heliothis v
20	382.5	16.6	515 5 Q46133	Q46133 locusta mig
21	382	16.5	552 5 P91765	P91765 myzus persi
22	380	16.4	484 11 Q70213	Q70213 cavia porce
23	379	16.4	568 5 Q9NFR5	Q9NFR5 drosophila
24	378.5	16.4	510 4 Q60854	Q60854 homo sapien
25	378	16.4	502 11 Q9JHD6	Q9JHD6 mus musculu
26	377	16.3	490 11 Q70212	Q70212 cavia porce
27	377	16.3	505 4 Q9BQ93	Q9BQ93 homo sapien
28	374.5	16.2	542 5 Q18556	Q18556 caenorhabdi
29	374.5	16.2	776 5 Q44202	Q44202 drosophila
30	372	16.1	502 4 Q96RH4	Q96RH4 homo sapien
31	371.5	16.1	489 4 Q9BRH2	Q9BRH2 homo sapien
32	371.5	16.1	531 5 Q96632	Q96632 heliothis v
33	370.5	16.0	489 4 Q96RH3	Q96RH3 homo sapien
34	370	16.0	483 6 Q9N0F4	Q9N0F4 mustela put
35	368	15.9	515 5 Q16926	Q16926 caenorhabdi
36	367	15.9	484 13 Q9PMS8	Q9PMS8 gallus gall
37	365.5	15.8	595 5 P91764	P91764 myzus persi
38	365	15.8	580 5 Q9VC72	Q9VC72 drosophila
39	364	15.8	477 11 Q62999	Q62999 rattus norv
40	362	15.7	489 11 Q61225	Q61225 mus musculu
41	359.5	15.6	432 11 Q90XK6	Q90XK6 mus musculu
42	358.5	15.5	772 5 Q9W3G6	Q9W3G6 drosophila
43	357.5	15.5	560 5 Q62083	Q62083 caenorhabdi
44	356	15.4	532 5 Q9U940	Q9U940 myzus persi
45	356	15.4	573 5 Q9GQ09	Q9GQ09 caenorhabdi

## ALIGNMENTS

RESULT	ID	Q9VP08	PRELIMINARY:	PRT:	441 AA.
AC	Q9VP08	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	CG11822 PROTEIN (NICOTINIC ACETYLCHOLINE RECEPTOR BETA 3 (DBETA3))				
DE	SUBUNIT (NICOTINIC ACETYLCHOLINE RECEPTOR BETA SUBUNIT).				
GN	CG11822 OR NACR2BETA-21C.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
KC	STRAIN=BERKELEY;				
KX	MEDLINE=20196006; PubMed=10731132;				
RA	Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yeandle M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers J., Blazey R.G., Champe M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,				
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Besono K.Y., Berens P.V., Bertman B.P., Bhandari D., Bolshakov S.,				
RA	Botkova D., Botchan M.R., Bouck J., Brockner P., Brothier P.,				
RA	Buttis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablos J.M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,				
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,				
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,				
RA	Jatelli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,				



RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SALIVARY GLAND;  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC011490.1; -  
 KW Receptor.  
 SQ SEQUENCE 512 AA; 58735 MW; AC0C6CC398B89C6C CRC64;

Query Match 18.8%; Score 434; DB 11; Length 512;  
 Best local Similarity 25.0%; Pred. No. 3.7e-31;  
 Matches 130; Conservative 91; Mismatches 179; Indels 120; Gaps 19;

QY 9 APVSGPGLPDLQMLMGLMGLTSGATATADPKNANKALDRHAGLFTNYDSVQ 68  
 DB 2 AP-SHAFOPWILYLMCLDLP--MPAVLA--QCGSHHADRDLFKHLFGGYNMAR 53  
 QY 69 VFGPTPTNVSLEMYVY-----IDIDELNGKLTTHCMNLNRDEDERVQPSQYDNIT 121  
 DB 54 V-----PNTSDVYVAFGLSIAGLIDVDEKNQMTTNVMLKQEMNDYKLRMDPAEFGNIT 108  
 QY 122 QITLKSEVWTPQITLFGNDEGLMAETGYLTH---DGHFRMPPAVYATACELMLMW 178  
 DB 109 SLKVPSEMIIDIVLYNNADGE-FAVTHMTKAHLFTGTVHWVPPAIYKSSCSIDVTF 167  
 QY 179 PHDKOSCKLKIGSW--GLKVLVPENGTAARGESLDHDLVQSPMEIYVSRHAFVSODV- 234  
 DB 168 PFDQCNCKKFKGSMYDRAKIDLEQ---MERTVDLKDWESEGMALINATGYNSKKTID 223  
 QY 235 -YGYMEYTLTAQRSSMTAVIYTPASCIIVILASFAFLPPHMGKEKIMINGLLII 289  
 DB 224 CCAEIPDVTYFVIRLPLFTINLIIPCLLISCLTVLFTVLPSC-GEKITLCSVLL 282  
 QY 290 VIAAFMTAQLLPVLSNTPLVITYSTLSLYSTVEVLVLYLATGKRR-----R 344  
 DB 283 SLTFVLLLTETIPISTSLVPLIGEXLFTMTFTLSIVTVFVLNV---HHRSPSTHN 338  
 QY 345 LPEALRKLHGHGTWLLS-----VFSTG-----ESQAEKREKMDENHY 385  
 DB 339 MPMWVAVALLGVRPMRLMNNRPLRPMELHSGPQLKLSPTYHMLTMDADERETEER 398  
 QY 386 EEADEQ-----ESSPLGI-----NHTEVPGAKAN----- 409  
 DB 399 EEDENICACAGLPOSSMGVLVGHGSLHRLANGPEAKTPSQASEILLSPQIQALEGVHY 458  
 QY 410 -----QPDALLATAVDRISVFSLAFLIL 435  
 DB 459 IADHLRSEDADSSVKEDMKRYAMVVDRI---FLMLFIIV 494

RESULT 4  
 Q9XS62 PRELIMINARY: PRT; 457 AA.

AC Q9XS62; 01-NOV-1999 (Tremblrel. 12, Created)  
 DT 01-NOV-1999 (Tremblrel. 12, last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA-SUBUNIT.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=MUSCLE;  
 RA MEDLINE=20092256; PubMed=10628675;  
 RA Yoshioka T., Uzuka Y., Tanabe S., Sarashina T., Ishiguro N.;  
 RT "Molecular cloning of the canine nicotinic acetylcholine receptor  
 alpha-subunit gene and development of the ELISA method to diagnose  
 myasthenia gravis";  
 RL Vet. Immunol. Immunopathol. 72:315-324(1999).  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 DR EMBL: AB021708; BAA76876.1; -  
 DR InterPro: IPR000188; GABA\_receptor.  
 DR InterPro: IPR001175; Neur\_channel.  
 DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
 DR Pfam: PF02932; Neur\_chan\_memb; 1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR PROSITE: PS00236; NEURON\_ION\_CHANNEL; 1.  
 KW Glycoprotein; Ionic channel; Postsynaptic membrane; Transmembrane.  
 SQ SEQUENCE 457 AA; 51878 MW; F71AD19AE2ECA3C8 CRC64;

Query Match 18.1%; Score 418.5; DB 6; Length 457;  
 Best local Similarity 26.6%; Pred. No. 8.3e-30;  
 Matches 123; Conservative 94; Mismatches 164; Indels 81; Gaps 21;

QY 26 MLMLGTSVPGATATADPKNANKALDRHAGLFTNYDSVQV---FQGTPTNVSLEMY 82  
 DB 7 LLLGLCSA-GVYVSGSEHT-----RLVAKLFEDYNSVVRVEDHQAVENTVGLQ- 57  
 QY 83 VYIIDDELNGKLTTHCMNLNRDEDERVQPSQYDNITQITLKSEVWTPQITLFGNDE 142  
 DB 58 IQLINDEVNQIYTTNVRKQGVVDNLRKNPEYDYGVRKIHIPSEKIRRPDLVLYNNAD 117  
 QY 143 G-GLMAETQVTLSDHGHFRMPPAVYATACELMLNMPHDKOSCKLKIGSW---GLKV 196  
 DB 118 GDAIYKFKVLLDYGTHTTPRALFKSYCELYVHPPEDQCNCKKLGITWYDOSVYA 177  
 QY 197 VLPENGTAARGESLDHDLVQSPMEIYVSRHAFVSODYG-----YME--YTTLAQR 246  
 DB 178 INDES-----DQPDLSNFMESEGMVIKESRGKHWV---FYACCPSTPYLDITYHVMQR 229  
 QY 247 RSMYNAVYTPASCIIVILASFAFLPPHMGKEKIMINGLLIIVIAAFMTAQLLPVLS 306  
 DB 230 LPIYFTVANNIIPCLFSEFTGLVFLPTD-SGKMKPLSISVLSLTFVLLVLYELIPST 288  
 QY 307 NNPPLVIVYSTLSLYSTIVEVLVLYATGKRR-----LPEALRKLHGHGTWL 361  
 DB 289 SAVPLIGKMLFTMNVAVIAIITVIV---NTHHRSPSTHMPWVKVF---IDTITP 341  
 QY 362 LLSVFSTGESEAEKTKEMDENHY-EEADEQESS-----PLG-----INHTEVPGA--- 406  
 DB 342 NIMFSTMKRPSREK---QDKRIFTEIDIDISDGKPPPMGFHSPLKHPVKSAIIG 398  
 QY 407 -----KANQ-----FDWALLATAVDRISVFSLAFLI 434  
 DB 399 VYIATETKSDQESNNAEMRYAMVMDHILLGVMLVCI 440

RESULT 5  
 Q918C7 PRELIMINARY: PRT; 452 AA.

AC Q918C7; 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 GN ALPHA 10 SUBUNIT OF NICOTINIC ACETYLCHOLINE RECEPTOR PRECURSOR.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianine;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SPLEEN;  
 RA Barabino B., Boulter J., Alliod C., Ballivet M.;  
 RT "Alpha 10 NAChR expression in the developing chick embryo";  
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 DR EMBL: AJ295624; CAB94482.1; -  
 DR InterPro: IPR000188; GABAA\_receptor.

DR InterPro: IPRO01175; Neur\_channel.  
 DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
 DR Pfam: PF02932; Neur\_chan\_memb; 1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; 1.  
 KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor; Signal;  
 FT Transmembrane. 1  
 FT SIGNAL 1 26 POTENTIAL.  
 FT CHAIN 27 452 ALPHA 10 SUBUNIT OF NICOTINIC  
 FT ACETYLCHOLINE RECEPTOR.  
 SQ SEQUENCE 452 AA; 50052 MW; 3836C31530BAAD8 CRC64;

Query Match 18.0%; Score 417; DB 13; Length 452;  
 Best Local Similarity 28.4%; Pred. No. 1.1e-29;  
 Matches 129; Conservative 73; Mismatches 190; Indels 62; Gaps 18;

QY 23 LMGMLGLTSPGATATADPKNANKALDRHAGLFTNYSDVPVQGT--PTNVSLE 80  
 DB 10 LLAISLSTGLVAPGCGA-----AQRHAKRLHDLRFANYSALRPA-EDTERALNVTIQ 62  
 QY 81 MVTYIDIDELNGKLTTHCWLNLNRDEERWQPSQYDNTITQITLKSEVWPTQITLF- 138  
 DB 63 VTLQIIMDERNOVLISLWVROAMLDALHMDKDAAGIDSTRIPSSYWRPDIVLYN 122  
 QY 139 NGDE--GGLMAETQVTLSDGHFRMPRPVYATVCELMLNMPHDQSKLKIGSWGKLV 196  
 DB 123 NADRFEGGSM-ETWVVLRSDDGHIMWDSCKVDVSYFPGDQCRLTFGSM---- 177  
 QY 197 VLPENGTAAGESL-----DHDDLVSPEWEIVDSRA-----HVSQDIYGYMYET 241  
 DB 178 -----TYNQIIDLRLNLDLTDGLDFEYENWEVLGNPAIRNVVYGCCEPPDYVT 231  
 QY 242 LTAQRSSMYAVIYTPASCIIVIALSAFWLPPHMGKEKIMINGLLIVIAFLMYFAQL 301  
 DB 232 LLERRASFYFNLLPCVMSLAPLGFYLPAD-SGKVSGLGVTLALVYFQLVAES 290  
 QY 302 LPLVSNPTLVVIFYSTLSLVYSTIVEVLVYL-ATGKHKRLRLPRLKRLHGLGTW 360  
 DB 291 MP-SESVPLIGKYYIAWMTITASTALTIFIMNHCGPGAPRPVPAWRLLIHLARA 349  
 QY 361 L-LLSVFSTGESQAEKTKEM-DEHPYEADQESSP---LGINHTEVPG----- 405  
 DB 350 LCVEVEGSGRPGREGGMPDPPEGVPGICLPSRCLCHHVAIVSGYIAGVER 409  
 QY 406 ---AKANQFDALLATVADR-ISVFSFLAFLI 434  
 DB 410 RHRTAQRRAAEMKKVAKYMDRFEMVPEFLMYFLM 443

RESULT 6  
 Q9ERT51 PRELIMINARY; PRT; 629 AA.  
 AC Q9ERT51  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 4 SUBUNIT.  
 GN CHRNA4 OR ACRA4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LONG SLEEP SELECTED LINE;  
 RA Stitzel J.A., Jimenez M., Smolen A., Modir J.;  
 RT "Cloning of mouse nicotinic acetylcholine receptor subunit alpha 4  
 RT cDNA";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 DR EMBL: AF225912; AAF34716.2; -.

DR MGD; MG1:87888; Chrna4.  
 DR InterPro: IPRO01175; Neur\_channel.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; 1.  
 KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;  
 FT Transmembrane.  
 SQ SEQUENCE 629 AA; 70319 MW; 78C28DEDFE371FB2F CRC64;

Query Match 18.0%; Score 417; DB 11; Length 629;  
 Best Local Similarity 28.6%; Pred. No. 1.8e-29;  
 Matches 98; Conservative 74; Mismatches 137; Indels 34; Gaps 9;

QY 11 VSGGRLPILQMLGMLGLTSPGATATADPKNANKALDRHAGLFTNYSDVPVQ 70  
 DB 3 IGGGAPPLLLILLILLILGLTPASSHETRAHAERLLKR---LFSGYNWSRPV- 57  
 QY 71 QGTPTNVSLEWVYV-----IDDELNGKLTTHCWLNLNRDEERWQPSQYDNTITQI 123  
 DB 58 ---ANISDVVLVRRGSLIAQLIDVDEKNQMTTNVWVKQEMHDKLRMDGDEVNTSI 113  
 QY 124 TLKSEVWPTQITLFNGDEGLMAETQVTLSH---DGHFRMPRPVYATVCELMLNMPH 180  
 DB 114 RIPSBLIRPDIVLYXNADGN-FAVTHLTRKALTYDGVQMTPPAIYKSSCIDVTFPPF 172  
 QY 181 DKQSKLKIGSW---GLKVLPENGTAGESLDHDDLVSPEWEIVDSRAHVSQDY-- 234  
 DB 173 DOONCTMKFGSMVYDKAKIDL---VSMHSRVQDLDFEWSGEWIVDAVGTYNRRKYEC 228  
 QY 235 ---YGYMYETLTAQRSSMTAVIYTPASCIIVIALSAFWLPPHMGKEKIMINGLLI 291  
 DB 229 AEIYPDITVAPRIKRLPLFYLINILIPCLLISCLTVLFIYFISBC-DEKVTLCISVLL 287  
 QY 292 AAFMYFAQLPLVSNPTLVVIFYSTLSLVYSTIVEVLVYL 334  
 DB 288 TVFLILLITEIIPSTSLVPLIGEVLLFTMTVTSIYTVVYL 330

RESULT 7  
 Q923N8 PRELIMINARY; PRT; 629 AA.  
 AC Q923N8  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE BM40117.2.1 (CHOLINERGIC RECEPTOR, NICOTINIC, ALPHA POLYPEPTIDE  
 DE 4 (ISOFORM 1)).  
 GN CHRNA4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wall M.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL450341; CAC40724.1; -.  
 KW Receptor.  
 SQ SEQUENCE 629 AA; 70304 MW; 261455B6ED50B4IC CRC64;

Query Match 17.8%; Score 412; DB 11; Length 629;  
 Best Local Similarity 28.6%; Pred. No. 5.2e-29;  
 Matches 98; Conservative 73; Mismatches 138; Indels 34; Gaps 9;

QY 11 VSGGRLPILQMLGMLGLTSPGATATADPKNANKALDRHAGLFTNYSDVPVQ 70  
 DB 3 IGGGAPPLLLILLILLILGLTPASSHETRAHAERLLKR---LFSGYNWSRPV- 57  
 QY 71 QGTPTNVSLEWVYV-----IDDELNGKLTTHCWLNLNRDEERWQPSQYDNTITQI 123  
 DB 58 ---ANISDVVLVRRGSLIAQLIDVDEKNQMTTNVWVKQEMHDKLRMDGDEVNTSI 113

QY	124	TLKSEWVTPDITLTFNADDEGLMAYETQVLTISH---DGHFRMPAPVAYTATCEINMLNMPH 180
Db	114	RIPSELIMRPDIYLYKXNADG-FAYTHLTKRAHLYEDRQVMTTPPAIKKSCSIDVIYFFP 172
QY	181	DKQSCCKLKIGSW---GLKVVLPENGSTARGBESLDHDDLVSQPEWEIVDSRAHFVSQDY-- 234
Db	173	DQQNCMTKMGFSWTFYDKAKIDL---VSMHSKRVDDLDPEWESGEWIVYDAVGYTMRKECC 228
QY	235	---GGMWEYTLTAQRBSMTAVIYTFPASCIVIALASAPLPRHMGCKIMINGLIIVY 291
Db	229	AEIYPDITVAFITIRLPLEPTINILIPCLLSICTLYVETFLPSEB-GEKVTLCISVLTSL 287
QY	292	AAFLMYFAQLLPVLNNTPLVIVFYSTSLLYSVSTIVEYLV 334
Db	288	TVFLLLITETIPSTSLVPIPLGELTPTMIFVLTSLIYIVFVL 330
RESULT	8	
Q923N7		
AC	Q923N7	PRELIMINARY; PRT; 641 AA.
DT	01-DEC-2001	(TREMBLrel. 19, Created)
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)
DE	B4M01117.2.2	(CHOLINERGIC RECEPTOR, NICOTINIC, ALPHA POLYPEPTIDE 4 (ISOFORM 2)).
GN	CHRNA4	
OS	Mus musculus (Mouse).	
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	
OX	NCBI_Taxid:10096;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Wall M.	
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AL150341; CAC40725.1; -.	
KW	Receptor.	
Q0	SEQUENCE 641 AA; 71747 MW; BEEBF1164C1C4A0 CRC64;	

	Query Match	17.8%	Score 412;	DB 11;	Length 641;	
	Best Local Similarity	28.6%;	Pred. No. 5.4e-29;			
	Matches 98; Conservative	73;	Mismatches 136;	Indels	34;	Gaps
OY	11 VSGPGLPLLQMLMGMLMGLTVPATATADPKNANKALDRLAGLTNTYDSDPVDF	70				
Dd	3 IGGSGAPPLLLLPILLILGTLLTPASSHIEFRAAEFLRLR----	LFSSGNKMSRP-	57			
OY	71 OGTFPNVSELMVVY-----IDLDLNGKITTHCWLNRDRDERWQPPOYLNIQI	123				
Dd	58 ----ANISDVVLVRGSLTAOLLDIVDEKKOMMTNVWAKOEHNDRKLRRDPPDYENVTSI	113				
OY	124 TLKSSEVWTPOITTEFNCGEGGLMAETOVYLSH---DGHRMPRPAYVTAYCELMNLMPB	180				
Dd	114 RIPSLLIRPDIVLVNNADGD-FAVTLHLTKALLFDGRQVMRPRAIKYSCSIDVFPEFF	172				
OY	181 DKOSKLIKIGSH---GLAVLPENGTANGESLDHDDLVQSFPWEIYDSRAHFVSODY---	234				
Dd	173 DOONCTMKFGSWTYKAIDT----VSMHSRDOLDPFWSGEMVYDAAGVTNTRKEYECC	228				
OY	235 ---YGMEYETTLAORRSSMYTAVITYTPASCIYIALSAFWLRPHMGGEKIMINGLIIVI	291				
Dd	229 AEIVPDIVAFRIIRLPLEFYFINLIIIPCLISICLVLYVEYLRSEC-GERVYTLCISVLSL	287				
OY	292 AAFLMYFAQLLPVLSNNPTLVVIEFYSSTLLYSVSTIVEVLV	334				
Dd	288 TWELLITEIIPSTSLVIPRLIGEYELFTWIEFTLSIVITVEVL	330				
RESULT	9					
C9N5U8						
ID	C9N5U8	PRELIMINARY;	PRT,	664 AA.		
KC	C9N5U8;					

DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE HYPOTHELICAL 75.8 KDA PROTEIN.  
 GN F27B3.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peioderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX Wilson R.;  
 RA "The sequence of C. elegans cosmid F27B3.";  
 RT Submitted (Mar-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX Waterston R.;  
 RA "Direct Submission.";  
 RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 RL -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 DR EMBL; AC006631; AAF39790.1; -;  
 DR InterPro: IPR001175; Neur channel.  
 DR PRINTS; PR00252; NRIONCHANNEL.  
 DR PROSITE; PS00236; NEUROTR\_ION\_CHANNEL; 1.  
 KW Glycoprotein; Hypothetical protein; Ionic channel;  
 KW Postsynaptic membrane; Transmembrane.  
 SO SEQUENCE 664 AA; 75804 MW; 636288145947CFA1 CRC64;

	Query Match	17.8%	Score 411.5	DB 5	Length 664
	Best Local Similarity	27.3%	Pred. No. 6.3e-29		
	Matches 104	Conservative	81	Mismatches 167	Indels 29
					Gaps 10
QY	27	LLMGITLSP-----GATATADPKRANKYALDRLAGLFTNYNDSDVOP-VFOGFTPNSL	79		
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :			
DB	55	IYGGIPSPISYKNGKANGAGNELSRADQNV--RLVRDLLDYDNNREPRSVHSKEPIINTVF	111		
QY	80	EMAVY-YIDIDELNGKLTHCHLNRMRDEEKYQWOSYDNTQITLKSEWTPQITLF	138		
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :			
DB	112	VFSLQIIDVDNRNLIITNSMIRLHWVDYKLVMPBRQLQVNTRIHIPSDKIMKPDILLY	171		
QY	139	NDEDEGLM---AEQVTLTSHDGHFRMPRPAYVTAACELNMLNMPHDQSKCLIKGSM---	192		
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :			
DB	172	NNAADQYKXSWSTDIYDYLGNHIMPLSAITTSQCPDLVKRTHPRDQICILKYSMAWD	231		
QY	193	GKLVLLPENGARGESLDHDLVQSPMEWIVDSRA-----HFVSODYYGYMEYTLTFAOR	246		
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :			
DB	232	GKKIDL---LTKSEQDGLTNYITTEWTESLGIRAEKNQVIYSCSEPRPFIDVHTIER	287		
QY	247	RSSMTAVIYTPASCVILALSAFNPRIIMGEGKIMINGLLIIVLAAPFLMYFAOLLPVLS	306		
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :			
DB	288	RMFVFVENLILPCVLISLALMGFTMPD-SEKAYILGTLISLSTVFVLMAYAEGRPPA	346		
QY	307	NNTPPLVIFYEYSLSLYLSTVSTVEVLYVL-ATGKHKKRLPEALKRLHLGHIGTWMLLSY	365		
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :			
DB	347	EALPLIGIYEGVTIMLALGTAMTGFVTVNIHTGVHGVRPFLQIFAPRYLSKILFVRI	406		
QY	366	FSTGESQAEKTKEMDENHYE	386		
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :			
DB	407	EPYHSIAHHVRYMYQKEHETE	427		

RESULT 10  
Q9BS29 PRELIMINARY; PRT; 484 AA.  
AC Q9BS29;  
DT 01-JUN-2001 (TReMBLrel. 17, Created)  
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE HYPOTHEICAL 55.9 KDA PROTEIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LONG CARCINOMA;  
RA Strausberg R.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
DR EMBL; BC004453; AA04453.1; -  
DR Interpro; IPR001005; Myb\_DNA\_bind.  
DR Interpro; IPR001175; Neur\_channel.  
DR PRINTS; PR00252; NRIONCHANNEL.  
DR PROSITE; PS00334; MYB\_2; UNKNOWN\_1.  
DR PROSITE; PS00236; NEUROTR\_ION\_CHANNEL; 1.  
KW Glycoprotein; Hypothetical protein; Ionic channel;  
KW Postsynaptic membrane; Transmembrane.  
SQ SEQUENCE 484 AA; 55894 MW; 74F70567007BCB51 CRC64;

Query Match 17.6%; Score 406; DB 4; Length 484;  
Best Local Similarity 26.0%; Pred. No. 1.3e-28;  
Matches 125; Conservative 84; Mismatches 201; Indels 70; Gaps 17;

QY 16 LPLLLMGLMGLTGLTSPGATATADPKNAVVALDRHLHGLFTNVDSDYQVFG-GTP 74  
DB 5 LAMLLMVOQALLLPLTLLAAGARSRRRTTRPALRLSLYLTRKGYRPRDMRKP 64  
QY 75 TNSLEMYV-TYIDDELNGKLTTHCWLNLWRDERVQSPQYDNTITLKSEVWTP 133  
DB 65 TVVSIDVIVAILNVDENQVLTYYIMYRQYWDDEFLQNMEDPDNTIKLSIPDSINVP 124  
QY 134 QITLFNG--DEGLMAETQVTLSDGHFRMPRAVYTAACGLNMLNPHDKOSCKLIGS 191  
DB 125 DI-LINEFVDVGKSPNIPYVIRHQGEVQNYKPLQVYACSLDIDYNEPFDVQNSLFTS 183  
QY 192 MGLKVV-----LPENGARGESLDHDDLVOSPMEIYDSRAHF-----VSODYGY 237  
DB 184 W-LHTIDINISLRLEBKYS-----DRSVFMNGEHELGLVLPFRFESMESSNYAE 237  
QY 238 MEYTLTAQRSSMYTAVIYTPASCIYIALSAFWLPPHMGGEKIMINGLLIYIAFLMY 297  
DB 238 MKEFYVIRRRPLFYVSSLLPSIFLWMDIVGYLPPN-SGERVSFKITLLGLYSVFLII 296  
QY 298 FAQLLPVLSNTPLVIFYSISLYSVTIVEVLVLYLNGKHKRL-----PEALKRL 353  
DB 297 VSDLPATAGTPLIGYFVVCMLLVISLAETIFIVLV-----HKODIQDPVPAMLRLV 353  
QY 354 HGLHGLTWLLSVFTGE-----SOAEKTK-----MDEH-----PYEADDEQESSP- 395  
DB 354 LERT-ANULLCIRESSTSORPATKTDSCSMNGHCSMGGRPODFEKPFRRCSPFP 412  
QY 396 -----LGINTEVVGAKANQF-----DWALLATAVDRISFVSFLAFLIAI 437  
DB 413 PPREASLAVGCLLQELSSIRQLEKREIREVARDWLRVGSVLDKLLFHIYLLAVLSI 472

RESULT 11  
Q9JUB5 PRELIMINARY; PRT; 447 AA.  
AC Q9JUB5;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR SUBUNIT.  
GN ALPHA10.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY;  
RX MEDLINE=21145873; PubMed=11248107;  
RA Elgoyhen A.B., Vetter D.E., Katz E., Rothlin C.V., Heinemann S.F.,  
RA Boulter J.;  
RT "alpha 10: A determinant of nicotinic cholinergic receptor function in  
RT mammalian vesicular and cochlear mechanosensory hair cells."  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3501-3506(2001).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
DR EMBL; AF196344; AAF27624.1; -  
DR Interpro; IPR001175; Neur\_channel.  
DR PRINTS; PR00252; NRIONCHANNEL.  
DR PROSITE; PS00236; NEUROTR\_ION\_CHANNEL; 1.  
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;  
KW Transmembrane.  
SQ SEQUENCE 447 AA; 49819 MW; EEE49D93490B696F CRC64;

Query Match 17.5%; Score 404.5; DB 11; Length 447;  
Best Local Similarity 27.7%; Pred. No. 1.5e-28;  
Matches 125; Conservative 76; Mismatches 191; Indels 57; Gaps 17;

QY 24 MGLMLGITSVPGATATADPKNAVVALDRHLHGLFTNVDSDYQVFGPT-NVSEMY 82  
DB 10 LGFLL--LFLPACLGAEGRLAH-----KLFRLDFAVYTSALRPVADDTQTLNWTLEV 62  
QY 83 VT-YIIDDELNGKLTTHCWLNLWRDERVQSPQYDNTITLKSEVWTPQTLTN-- 139  
DB 63 LSGIIDLDERQVLTLYLMIRQEWTDAYLHMDPRAYGLDAIRIPSLRWMPDITVLYNKA 122  
QY 140 GDEGLMAETQVTLSDGHFRMPRAVYTAACGLNMLNPHDKOSCKLIGSM---GLKV 196  
DB 123 DTQPPASSTNVVVRHDAVMDAPAITRSSCRDVSAPFPDADRCGLTESWTHGHOL 182  
QY 197 -VLEPENGARGESLDHDDLVOSPMEIYDSRA-----HFVSODYGYMEYTLTAQRSS 249  
DB 183 DVPR-----RGTSASLADFEVNEVRVYLGMPARRRVLTGYGCCSPDPVPTLLRRRA 237  
QY 250 MYTAVIYTPASCIYIALSAFWLPPHMGGEKIMINGLLIYIAFLMYFQALLVLSNNT 309  
DB 238 AYVCNLLPVCVLSILAPLAHELPAD-SGEKVSIGYVLLALVTFQILLASMP-RESV 295  
QY 310 PLVVFYSTSLVSVTIVEVLVYL-ATGKHKRLPEALKRLHGLHGLTWLLSVFST 368  
DB 296 PLIKYVMAITMTVTFSTALITILIMNLHYCGPNHPPAPAKRVLLGLHLAGLEV---RE 352  
QY 369 TGESQAEKTEKMDHPYEADDEQESSPLG-----INHTEV-----PGA 406  
DB 353 RGE-PCGSKPLESAP--SLQPPASAPGCPHEBRCLCHQALLHHTIASFTRSHRA 409  
QY 407 KANQDMALLATAVDRISFVSFLAFLIAI 437  
DB 410 QRHEDWKRRLAVMDREFLGIFFCMALVMSL 440

RESULT 12  
Q9ROW9 PRELIMINARY; PRT; 494 AA.  
AC Q9ROW9;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE NICOTINIC ACETYLCHOLINE RECEPTOR SUBUNIT ALPHA6.  
GN CHRN6 OR NTC6.





"Preliminary: Cloning of nicotinic acetylcholine receptor subunits of *Rt* *Heliothis virescens*.";  
 RT Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC - SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 DR EMBL: AJ000399; CA004056.1; -;  
 DR InterPro: IPR001188; GABAA\_receptor.  
 DR InterPro: IPR001175; Neur\_chan.  
 DR Pfam: PF02931; Neur\_chan\_LBD.1.  
 DR Pfam: PF02932; Neur\_chan\_memb.1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR PROSITE: PS00236; NEURCTR\_ION\_CHANNEL.1.  
 KM Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor; Signal;  
 KW Transmembrane.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 523 NICOTINIC ACETYLCHOLINE RECEPTOR.  
 SQ SEQUENCE 523 AA; 59214 MW; 7E39A035B6BA00FD CRC64;

Query Match 17.1%; Score 395; DB 5; Length 523;  
 Best Local Similarity 23.7%; Pred. No. 1.4e-27;  
 Matches 115; Conservative 91; Mismatches 188; Indels 92; Gaps 13;

QY 26 MLMGLTSGCATATAPKNNANVALDRLHAGLFTNYDSQVPFGQIPTVSLDMVY-- 83  
 DB 14 LLVLAALAGCAAMPK-----RLYDILLNKNKLVPLV-----NVSALTVRI 60  
 QY 84 ----YIDIDELNGKLTTHCMLNLRWDEERWQPSQYDNIQTITLKSEVMTPTLF 138  
 DB 61 KLIKSLQIDVNLKQIMTNTLMVQSWDYKLSWEPREYGVEMLHVPDSHIMRPDIYLY 120  
 QY 139 NGDEGG--LMAEQVTLTSHDGHFRMPPAVYAYCELMNLMPHDQSKLKIGSW--G 193  
 DB 121 NMDAGNEVTLATKATINTGTGVEWEPATYKSSCEIDVEYFPDQGVCMKFGSWTYDG 180  
 QY 194 LKVVLPENGARGES-----LDHDDLYQSPMEIYD-----SRAHVSQDYGYGMETL 242  
 DB 181 FQVDLRIHIDEKGTNVVELGDLSEFYTSVMDILEVPAVANEKFFYCCDDPYLDITFNI 240  
 QY 243 TAQRSSMYTAIVITPASCIVIALSAFWLPPHMGGEKIMINGLLIYIAFLWYFAQL 302  
 DB 241 TMRKRTLEYTNLITPCMGISFLVLYFPLPSD--SGEVSLSITLSITVFFLLAEIT 299  
 QY 303 PVLSNNTPLV--VIFYSTSLYLSVSTIVEVLYLYLATGKHRRLEALRRLHGLGTW 360  
 DB 300 PPTSLVPLDLCKEFLVLMIDTFESICVTVVVLANVHFRSPQHTMAVWVAVFIH-VLPRL 358  
 QY 361 LLL-----SVFSTGESQAETKEMDEHP----- 384  
 DB 359 LVMRRPHYRVDPHRSRPAGLVTAVESAPWEDGSPLAGLAGGPGPPACACRSCRLH 418  
 QY 385 -----YEEADQESSPLGINHTEVPGAKANOF-----DMALLATAVDRISFVSF 428  
 DB 419 DAPALCALRMHRCPELNKAIDGINIYAEQTRKEESTRVEKMKYVAMVLDLFLMITF 478  
 QY 429 SLAFLI 434  
 DB 479 TLAVVV 484

RESULT 15  
 Q96FB8 PRELIMINARY; PRT; 501 AA.  
 AC Q96FB8;  
 DT 01-DEC-2001 (TEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)  
 DE 01-DEC-2001 (TEMBLrel. 19, Last annotation update)  
 DE SIMILAR TO CHOLINERGIC RECEPTOR, NICOTINIC, BETA POLYPEPTIDE 1 (MUSCLE).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=EYE, AND RETINOBLASTOMA;  
 RA Strausberg R.;  
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC01371; AAH1371.1; -;  
 KW Receptor.  
 SQ SEQUENCE 501 AA; 56698 MW; 365CBFA795A51394 CRC64;

Query Match 17.0%; Score 392.5; DB 4; Length 501;  
 Best Local Similarity 22.6%; Pred. No. 2.3e-27;  
 Matches 117; Conservative 91; Mismatches 178; Indels 131; Gaps 17;

QY 17 PLLQIMLGMGLTSGCATATADPKNNANVALDRLHAGLFTNYDSQVPFGQ--GTPR 75  
 DB 3 PGALLMLGLL--GAPLAPGVRS-----EADGRRLRELFEGSYSSVAPAEVODRV 52  
 QY 76 NVSLDMVYT-YIDIDELNGKLTTHCMLNLRWDEERWQPSQYDNIQTITLKSEVMTPT 134  
 DB 53 RVSGLTIAQLISLNKDEEMSTKVLDELWTDYRLSWDPAHGDIDSLRTAESVWLPD 112  
 QY 135 ITLNGDEGG--LMAEQVTLTSHDGHFRMPPAVYAYCELMNLMPHDQSKLKIGSW 192  
 DB 113 VLLNNMDGNEDVALDISVAVSSDGSVWQPPGAYRSSCSIQVTFPFDWQNCITMVESSY 172  
 QY 193 -----GLKVVLPENGARGESLDHD-DLYQSPMEIYDSRAHFVSGQ--DYGYME-- 239  
 DB 173 SYDSSEVSLQGTGLPDQGHQEHHEGTFTENGOMELIHKPSLIQPRGPRGREGOR 232  
 QY 240 ----YTLTAQRSSMYTAIVITPASCIVIALSAFWLPPHMGGEKIMINGLLIYIAFL 295  
 DB 233 QEVIFYLIIRKRPFLFYLVNVAIPCLITLLAIFEVYLPD--AGEKMGISIFALLTLTVFL 291  
 QY 296 MYFAQLPLVLSNNTPLV--VIFYSTSLYLSVSTIVEVLYLYLATGKHRRLEALRRLH 354  
 DB 292 LLLADKVPETSLVPLITIKYMLFTVLTVEFVLSVAVLNHRSRPTHQPLWVROIPT 351  
 QY 355 GHLGTWLLLSVFTSTGESQAETKEMDEHPYEEADQESSPLGINHTEVPGA----- 406  
 DB 352 HKLPLYLRL-----KRPRKEDLMEPP-----HSSPSGSGRGRTDE 389  
 QY 407 -----KANOF----- 411  
 DB 390 YFIRKPPSDFLPKPNRFPPELAPDLRRFTDGNRAVALPELREVVSSISYAROLQE 449  
 QY 412 -----DMALLATAVDRI--SPVSF--SLAFLIL 435  
 DB 450 QEDHDALKEDMQFVAMVVDRLMTFTLETFSVGTLYI 486

Search completed: July 18, 2002, 10:29:21  
 Job time: 669 sec

Fri Jul 19 08:26:24 2002

us-09-732-680a-2.rspt

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 10:16:52 ; Search time 35.01 Seconds

(without alignments)  
487.727 Million cell updates/sec

Title: US-09-732-680A-2

Perfect score: 2311

Sequence: 1 MTTTPKIKAVSGPLPLL.....RISFVSPSLAFLILAIRCSV 441

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	444.5	19.2	528	1	ACH2_CHICK
2	427	18.5	457	1	ACHA_BOVIN
3	423	18.3	511	1	ACH2_RAT
4	418	18.1	627	1	ACH4_HUMAN
5	416	18.0	529	1	ACH2_HUMAN
6	414.5	17.9	457	1	ACHA_RAT
7	413.5	17.9	493	1	ACH6_RAT
8	412.5	17.8	457	1	ACHA_MOUSE
9	410	17.7	455	1	ACHA_CHICK
10	409	17.7	450	1	ACHX_HUMAN
11	407.5	17.6	622	1	ACH4_CHICK
12	407	17.6	557	1	ACH1_SCHER
13	407	17.6	630	1	ACH4_RAT
14	404.5	17.5	478	1	5HT3_HUMAN
15	403.5	17.5	464	1	ACHO_RAT
16	402.5	17.4	519	1	ACH4_DROME
17	400.5	17.3	457	1	ACH2_XENLA
18	398.5	17.2	457	1	ACH1_XENLA
19	397.5	17.2	468	1	ACH5_HUMAN
20	397	17.2	479	1	ACH9_RAT
21	397	17.2	505	1	ACHB_BOVIN
22	394	17.0	494	1	ACH6_HUMAN
23	393.5	17.0	479	1	ACH9_HUMAN
24	392	17.0	494	1	ACH6_CHICK
25	391.5	16.9	512	1	ACH3_CARAU
26	391	16.9	458	1	ACHO_HUMAN
27	391	16.9	493	1	ACH2_CAEL
28	390.5	16.9	456	1	ACHA_BRARE
29	390.5	16.9	456	1	ACHA_CHICK
30	390.5	16.9	507	1	ACH7_CAEL
31	388	16.8	482	1	ACHA_HUMAN
32	386.5	16.7	461	1	ACHA_TORMA
33	386.5	16.7	499	1	ACH3_RAT

## ALIGNMENTS

RESULT ID	ACH2_CHICK	STANDARD:	PRT:	528 AA.
AC	P09480;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-MAR-1989 (Rel. 10, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Neuronal acetylcholine receptor protein, alpha-2 chain precursor.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=88283624; PubMed=3267226;			
RA	Net P., Oneyser C., Allod C., Couturier S., Ballivet M.;			
RT	"Genes expressed in the brain define three distinct neuronal			
RT	nicotinic acetylcholine receptors.";			
RL	EMBO J. 7:595-601(1988).			
CC	- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN			
CC	EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND			
CC	LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA			
CC	MEMBRANE.			
CC	- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT			
CC	TYPE OF SUBUNITS: ALPHA AND NON-ALPHA (ALSO CALLED BETA). A			
CC	FUNCTIONAL RECEPTOR SEEMS TO CONSIST OF TWO ALPHA-CHAINS AND			
CC	THREE NON-ALPHA CHAINS.			
CC	- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			
CC	-----			
CC	EMBL: X07339; CAB59645.1; JOINED.			P09690 mus musculus
DR	EMBL: X07340; CAB59645.1; JOINED.			P11230 homo sapien
DR	EMBL: X07341; CAB59645.1; JOINED.			O93149 caenorhabdi
DR	EMBL: X07342; CAB59645.1; JOINED.			P02710 torpedo cal
DR	EMBL: X07343; CAB59645.1; JOINED.			P26152 gallus gal1
DR	EMBL: X07344; CAB59645.1; JOINED.			P49582 carassius a
DR	EMBL: AJ250360; CAB59625.1; JOINED.			P13908 carassius a
DR	PIR: S00377; ACH42N.			P36544 homo sapien
DR	InterPro: IPR000188; GABAA_receptor.			P25109 ratius norv
DR	InterPro: IPR001175; Neur_channel.			P02712 torpedo cal
DR	Pfam: PF02931; Neur_chan_LBD; 1.			P02713 gallus gal1
DR	Pfam: PF02932; Neur_chan_memb; 1.			
DR	PRINTS: PR00252; NRIONCHANNEL.			
DR	PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.			
DR	Postsynaptic membrane; Ionic channel; Glycoprotein; signal;			
KW	Transmembrane; Multigene family.			

FT SIGNAL 1 23  
 FT CHAIN 24 528  
 FT DOMAIN 24 239  
 FT TRANSMEM 240 264  
 FT TRANSMEM 272 290  
 FT TRANSMEM 306 327  
 FT DOMAIN 328 501  
 FT TRANSMEM 502 520  
 FT DOMAIN 394 410  
 FT DISULFID 158 172  
 FT DISULFID 222 223  
 FT CARBOHYD 54 54  
 FT CARBOHYD 104 104  
 SQ SEQUENCE 528 AA; 60675 MW; E76C6360AF876364 CRC64;

NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,  
 ALPHA-2 CHAIN.  
 EXTRACELLULAR.  
 CYTOPLASMIC.  
 POLY-GLU.  
 BY SIMILARITY.  
 ASSOCIATED WITH RECEPTOR ACTIVATION  
 (BY SIMILARITY).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 19.28; Score 444.5; DB 1; Length 528;  
 Best Local Similarity 27.7%; Pred. No. 3.1e-28;  
 Matches 119; Conservative 82; Mismatches 147; Indels 81; Gaps 15;

39 ATADPKNANKALDLRLAGLFTNVDSDVQPFQSTPTNVSLEMYVY-----IDDEL 91  
 22 ATRQKQPHGFADRLKHLFTGYNRMSRP-----PNTSDVIVYKFGSLAQILDVDEK 76  
 92 NGKLTTHCWLINLRDEERWQPSQYDNTQITLKSSEVWPTITLFGDEGLMAETQV 151  
 77 NQMMTIVWLKQEWSDYKLRNRPEDPNVTSIRVSEMIWIPDLVLYNNADGE-FAYTHM 135  
 152 TLST---DGEFRMPAPVATYACELNMLNMPHDKSKCKLIGSV---GLKVVLENGSTAR 205  
 136 TKALHFGNGKWKVPAYVSSCIDVTYFPFDQNCMKKFGSWTYDKAKIDL----- 188  
 206 GESLDH---DDLQVSEMEIYDSRAHFVSODY-----GYMYTTLAQRSSMTAVI 255  
 189 -ENNEHHVDLKDWESEGMALINAGRNRSKRYDCCTEIRPDITFYVIRLPLEYITNL 247  
 256 YTPASCIIVIALSAFWLPPHMGKIKIMINGLIIIVIAFLMAYFQQLPVLNNPLVAVIF 315  
 248 IIPOLLISCLVLFVFLPSPC-GEKITLCLISVLSLTVFLLITFEIIPISLVPLIGEX 306  
 316 YTSLSLYLSTVIVEVLVLYLANGKHKR-----RLPALRKLHLGLGTW-----LLL 363  
 307 LLEFMIVFTLSIITVEVLNV---HHRSPSTHMPHMSVFFGFIPLRMFKRPPDLL 362  
 364 SVFSTTG-----ESQAEKTKEMDEHYEADPOESSPLGINHTEV 403  
 363 PABSTTQYDPPTGRLSTSRQWLETDVDVQKWEDEEEEEEKAKVP-----SRV 417  
 404 P--GAKANQ 410  
 418 PSGSGSQCTQ 426

RESULT 2  
 ACHA\_BOVIN STANDARD; PRT; 457 AA.  
 AC P02709;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Acetylcholine receptor protein, alpha chain precursor.  
 GN CHRNA1.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 OC NCBI\_TaxID=9913;  
 RN NCBI\_TaxID=9913;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84039794; PubMed=6688857;  
 RA Noda M., Furutani Y., Takahashi H., Toyosato M., Tanabe T.,

RA Shimizu S., Kikuyotani S., Kayano T., Hirose T., Inayama S., Numa S.;  
 RT "Cloning and sequence analysis of calf cDNA and human genomic DNA  
 RT encoding alpha-subunit precursor of muscle acetylcholine receptor.";  
 RL Nature 305:818-823(1983).  
 CC -!- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN  
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND  
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA  
 CC MEMBRANE.  
 CC -!- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, AND ONE EACH OF THE BETA,  
 CC DELTA, AND GAMMA (IN IMMATURE MUSCLE) OR EPSILON (IN MATURE  
 CC MUSCLE) CHAINS.  
 CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X02509; CA26345.1; -.  
 DR PIR: A03169; ACBOAL.  
 DR InterPro: IPR000188; GABAA\_receptor.  
 DR InterPro: IPR001175; Neur\_channel.  
 DR Pfam: PF02931; Neur\_chan\_LBD. 1.  
 DR Pfam: PF02932; Neur\_chan\_membr. 1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL. 1.  
 KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
 KW Transmembrane.  
 FT SIGNAL 1 20  
 FT CHAIN 21 457  
 FT DOMAIN 21 230  
 FT TRANSMEM 231 255  
 FT TRANSMEM 263 281  
 FT TRANSMEM 297 316  
 FT DOMAIN 317 428  
 FT TRANSMEM 429 447  
 FT DISULFID 148 162  
 FT DISULFID 212 213  
 FT CARBOHYD 161 161  
 SQ SEQUENCE 457 AA; 51947 MW; 402613D35FD47A32 CRC64;

Query Match 18.58; Score 427; DB 1; Length 457;  
 Best Local Similarity 27.4%; Pred. No. 6.5e-27;  
 Matches 129; Conservative 91; Mismatches 163; Indels 88; Gaps 22;

17 PLLIOMLMGLMLGTVSPGATADPKNANKALDLRLHGLFTNVDSDVQV---FGGT 73  
 5 PLLL--LLGLCSAGL--VLGSEHET-----RLVAKLFEDINSVRPVEDNRQAV 49  
 74 PTNVSLMEVVTYIDIDELNGKLTTHCWLINLRDEERWQPSQYDNTQITLKSSEVWPT 133  
 50 EVTVGLQL-IQLINVDENVQITVTVNRLLKQWVDYNNKMPDDYGVKIKIHISEKIMRP 108  
 134 QITLFGNDEG--GLMAETQVTLSDHGFRRMPAPVATYACELNMLNMPHDKSKCKLIGS 191  
 109 DLVLYNNADDFALVKTFTKLLDYTGHTWTPPAIFKSYCEIIVTHRPFDONCSMKLGT 168  
 192 W---GLKVVLENGSTARSGSLDHDLDVQSPMEVIVSRA--HFVSQDYGG-----YME 239  
 169 WYIDGQSVVAINPES---DQPDLSNFMESGEVWIKESKGMKHW--FYACCPSTPYLD 220  
 240 --YTLTAQRSSMTAVIYTPASCIIVIALSAFWLPPHMGKIKIMINGLIIIVIAFLMY 297  
 221 ITHFVWQRLPLVFIYVAILPCLLFSFLGLVFLPLPD-SGKMTLSISLVSLTVFLV 279  
 298 FAQLLPVLSNTPLVVIFVYSTSLIYSVSTIVEVLVLYLANGKHKR-----LPEALRKL 352

```

Db 280 IVELIPSTSSAVPLIGKMYLFTMVFIASIIITVIYI---NTHHRSPTSHVMEVVKY 335
QY 353 LHGHLGTWLLLSVFTSGSAQKTKEMDEHPY-PEADQESS-----PLG-----INH 400
Db 336 F---IDTIPNIMFSTMKRPSREK---ODKLTFTEDIDISDTSKRGPPMKGFHSPLIK 389
QY 401 FEVPGA-----KANQ-----FDWLLATAVDRISFVSLAFLI 434
Db 390 PEVKAIEGKIYIAETMKSDESNNAAEEMKYVAMVMDHLLAVFALVCI 440

RESULT 3
ACH2_RAT
ID ACH2_RAT STANDARD: PRT: 511 AA.
AC P12389: 008952:
DT 01-OCT-1989 (Rel. 12, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-2 chain precursor.
GN CHRNA2 OR ACRA2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=88178113; PUBMED=2832952;
RA Wada K., Ballivet M., Boulter J., Connolly J.G., Wada E.,
RA Denaris E.S., Swanson L.W., Heinemann S.F., Patrick J.;
RT "Functional expression of a new pharmacological subtype of brain
RT nicotinic acetylcholine receptor.";
RL Science 240:330-334(1988).
[2]
RN REVISIONS.
RP STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RA Boulter J.;
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC MEMBRANE.
CC -1- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT TYPES
CC OF SUBUNITS: ALPHA AND NON-ALPHA (BETA). ALPHA-2 SUBUNIT CAN BE
CC COMBINED TO BETA-2 OR BETA-4 TO GIVE RISE TO FUNCTIONAL RECEPTORS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L10077; AAB60900.1; -
DR EMBL: M20297; AAA40664.1; -
DR EMBL: M20292; AAA40664.1; JOINED.
DR EMBL: M20293; AAA40664.1; JOINED.
DR EMBL: M20294; AAA40664.1; JOINED.
DR EMBL: M20295; AAA40664.1; JOINED.
DR EMBL: M20296; AAA40664.1; JOINED.
DR InterPro: IPR00188; GABA_A_receptor.
DR InterPro: IPR001175; Neur_chanell.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_NBD; 1.
DR PRINTS: PR00252; NRIONCHANNEL.
DR PROSITE: PS00336; NEURORRION_CHANNEL; 1.
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family.
FT SIGNAL 1 27
FT CHAIN 28 511 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,

```

```

FT DOMAIN 28 241 ALPHA-2 CHAIN.
FT TRANSMEM 242 266 EXTRACELLULAR.
FT TRANSMEM 274 292 POTENTIAL.
FT TRANSMEM 308 329 POTENTIAL.
FT DOMAIN 330 348 CYTOPLASMIC.
FT TRANSMEM 385 503 POTENTIAL.
FT DOMAIN 391 402 POLY-GID.
FT DISULFID 160 174 BY SIMILARITY.
FT DISULFID 224 225 ASSOCIATED WITH RECEPTOR ACTIVATION
FT CARBOHYD 56 56 (BY SIMILARITY).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 494 494 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 511 AA; 58611 MW; 3824E83BB01D613B CRC64;

Query Match 18.3%; Score 423; DB 1; Length 511;
Best Local Similarity 24.5%; Pred. No. 1,66-26;
Matches 126; Conservative 94; Mismatches 167; Indels 128; Gaps 18;

QY 20 LQMLGMLMLGLTSVPGATATADPKNANYKALDRHAGLFTYDSDVOPFQGTPTNVS 79
Db 8 LQFWTHLYLWCLLPV---AVLQGSHTHARDRLFKHLFGYNNRARPV-----PNTSD 59
QY 80 ENAVTY-----IDDELNGKLTHCWLNLNRDEERYWQSDNTITQILKSEVWT 132
Db 60 VVIVREGSLIAOLIDVENQMMTNNVWKQEMNPKYKLMWDAEFGNVTSLRVPSEMTWI 119
QY 133 POTTENGEDEGLMAYETVTLSH---DGHFRMPAPVAVYAGELMNLMPHDKCKIKI 189
Db 120 PDIYLYNNADGE-FAVTHTKALHFFGTGVHWPAPYIKSSSIDVTFPPDQCKMKF 178
QY 190 GSW---GLKVLDPENGATAGESLDHDDIVQSPMEIVDSRAHVSQDY-----YGYMEY 240
Db 179 GSWTYDKAKIDLEQ---MERTVDLKDWESEGMALINATGYNSKKYCCCAEIPDVY 234
QY 241 TLTQRRSMYAVITTPASCIIVIALSFMLPRHGGKIMINGILLIIVIAFLMFPQ 300
Db 235 YFVIRRLPLFYINILIPCLLISCLIVLFLPSECGEKIKLISVLSLTVFLLITE 293
QY 301 LPLVSNPTPLVIVFSTSLYLSVYIEVYLVLYATGKHKR-----RLPALKRLHG 355
Db 294 ITPSTSLVPLIGGYLFTMIVTISYITVVLNV---HHRSTHMPMVMVYALIG 349
QY 356 HLGTYLLS-----VESTGESQAERTKEMDEHPYEADQESS 393
Db 350 RVPRLMNMRLPRMELHGPSPLKISPYHMLETNNDACERETEE-----EEEEEDEN 403
QY 394 -----SPGGINH-----TEVPGAKANQF----- 411
Db 404 ICVCAGLPDSSMGVLYGHGGLHRLAMEPTKTP-SQASEILLSPQIOKALEGVHIADRL 462
QY 412 -----DWALLATAVDRISFVSLAFLI 435
Db 463 RSEDDSDSVKEDMKYVAMVVDRI---FLMLFTIV 493

RESULT 4
ACH4_HUMAN
ID ACH4_HUMAN STANDARD: PRT: 627 AA.
AC P43681:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-4 chain precursor.
GN CHRNA4 OR ACRA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=95237608; PubMed=7721089;  
 RA Monteggia L.M., Gopalakrishnan M., Touna E., Idler K.B., Nash N.,  
 RA Ament S.P., Sullivan J.P., Giordano T.;  
 RT "Cloning and transient expression of genes encoding the human alpha-4  
 RT subunits.";  
 RL Gene 155:189-193(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96430009; PubMed=8833159;  
 RA Steinlein O.K., Welland S., Stood J., Propping P.;  
 RT "Exon-intron structure of the human neuronal nicotinic acetylcholine  
 RT receptor alpha 4 subunit (CHRNA4).";  
 RL Genomics 32:289-294(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Hippocampus;  
 RX MEDLINE=97062879; PubMed=8906617;  
 RA Elliott K.J., Ellis S.B., Berckhan K.J., Urrutia A.,  
 RA Chavez-Noriega L.E., Johnson E.C., Veliceljevi G., Harpold M.M.;  
 RT "Comparative structure of human neuronal alpha 2-alpha 7 and beta  
 RT 2-beta 4 nicotinic acetylcholine receptor subunits and functional  
 RT expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and  
 RT beta 4 subunits.";  
 RL J. Mol. Neurosci. 7:217-228(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97162233; PubMed=9009220;  
 RA Groth Kormelink P.J., Luyten W.H.M.L.;  
 RT "Cloning and sequence of full-length cDNAs encoding the human neuronal  
 RT nicotinic acetylcholine receptor (nAChR) subunits beta3 and beta4 and  
 RT expression of seven nAChR subunits in the human neuroblastoma cell  
 RT line SH-SY5Y and/or IMR-32.";  
 RL FEBS Lett. 400:309-314(1997).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P., Clee C.M.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R., Corry N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lechvasi M.H., Leverish M., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S., McConachie L.J., McIlroy K., McMurray A.A.,  
 RA Oliver S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Ruce C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
 RA Stuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 RN [6]  
 RP SEQUENCE OF 26-627 FROM N.A.  
 RC TISSUE=Brain;  
 RA Mamalaki A., Remoundos M., Tzartos S.;  
 RL Submitted (May-1995) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP VARIANT ENFL PHE-280.  
 RX MEDLINE=96024658; PubMed=7550350;

RA Steinlein O.K., Mulley J.C., Propping P., Wallace R.H., Phillips H.A.,  
 RA Sutherland G.R., Scheffer I.E., Berkovic S.F.;  
 RT "A missense mutation in the neuronal nicotinic acetylcholine receptor  
 RT alpha-4 subunit is associated with autosomal dominant nocturnal  
 RT frontal lobe epilepsy.";  
 RL Nat. Genet. 11:201-203(1995).  
 CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE AChR RESPONDS BY AN  
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND  
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA  
 CC MEMBRANE.  
 CC -1- SUBUNIT: NEURONAL AChR SEEMS TO BE COMPOSED OF TWO DIFFERENT  
 CC TYPE OF SUBUNITS: ALPHA AND NON-ALPHA (BETA).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- DISEASE: Defects in CHRNA4 are the cause of autosomal dominant  
 CC nocturnal frontal lobe epilepsy type 1 (ENFL1 or ADNFLE). It is a  
 CC disease characterized by clusters of motor seizures during sleep.  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L35901; AAA64743.1; -;  
 DR EMBL: X89741; CAA61893.1; -;  
 DR EMBL: X89742; CAA61893.1; JOINED.  
 DR EMBL: X89743; CAA61893.1; JOINED.  
 DR EMBL: X89744; CAA61893.1; JOINED.  
 DR EMBL: X89745; CAA61893.1; JOINED.  
 DR EMBL: X89746; CAA61893.1; JOINED.  
 DR EMBL: U62433; AAB40111.1; -;  
 DR EMBL: X08421; CAA69698.1; -;  
 DR EMBL: AL121827; CAC36119.1; -;  
 DR EMBL: X87629; CAA60959.1; -;  
 DR MIM: 118504; -;  
 DR MIM: 600513; -;  
 DR InterPro: IPR000188; GABAA\_receptor.  
 DR InterPro: IPR001175; Neur\_channel.  
 DR Pfam: PF02931; Neur\_chan\_LBD.1.  
 DR Pfam: PF02932; Neur\_chan\_membr.1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL.1.  
 KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
 KW Transmembrane; Multigene family; Disease mutation.  
 FT SIGNAL 1 28  
 FT CHAIN 29 627  
 FT DOMAIN 29 242  
 FT TRANSMEM 243 267  
 FT TRANSMEM 275 293  
 FT TRANSMEM 309 330  
 FT DOMAIN 331 600  
 FT TRANSMEM 601 619  
 FT DISULFID 161 175  
 FT DISULFID 225 226  
 FT VARIANT 280 280  
 FT S->F (IN ENFL1).  
 FT /FTID=VAR\_000295.  
 FT SEQUENCE 627 AA; 69957 MW; B3A0C0151E5A2A8 CRC64;  
 SQ  
 Query Match 18 18; Score 418; DB 1; Length 627;  
 Best Local Similarity 29.88; Pred. No. 5.2e-26;  
 Matches 102; Conservative 67; Mismatches 139; Indels 34; Gaps 9;  
 QY 11 VSGGLPLQLQMLGMLCLTSPGATADPKNANVKAALDRHAGLFTNYSQVDF 70  
 Db 3 LGGPGARLRLLPRLLLGLTGTLRRSSHYET-----RAHAEERLKKLFEGYKWSRPV 55  
 QY 71 QGTPTNVSLSEAVYTY-----IDIDELNGKLTTHCWLNRWDERVWQPSQYDNTIQT 123



```

Db 56 ----ANISDVVLVRFGLSTIAQLIDYDEKQMMTINWVWQEHNDYKLRMDPADYEIVTSI 111
QY 124 TLKSEVWTPQTLFNGDEGGGLMAE--TQVYTSIHGCFHFMPPAYTATVCELMLNMPHD 181
Db 112 RIPELIMPRDIDLVLNNAAGDAVYTHLTRAKHLFHDGRVOMTPAIYKSSCIDYVEFFPPD 171
QY 182 KOSCKLKISW---GLKVLVPENGTARGESLDHDDLVSQPEMEIVDSRAHFVSODY---- 234
Db 172 QONCTIKKESWTYDKAKIDL-VNMSR---VDQDPWEGENVYIDAVCTITNTRKECA 227
QY 235 --YGYMEYTLFAORSSMYTAVIYTPASCIIVLALSAFMLPPHMGSEKIMINGLLITVIA 292
Db 228 ELYPDITVFAVIRLPLFYTIMLIPCLLISCLTVLVEFLPSPC-GEKITLICISVLSTLT 286
QY 293 AFLMFAQLLPVLSNNTPLVVFYTSLSLYSVSTIVEVLV 334
Db 287 VELLLETIELPSTSLVPLIGELYLFTMLFVLSIVITVEVL 328

RESULT 5
ACH2_HUMAN
ID ACH2_HUMAN STANDARD: PRT: 529 AA.
AC Q15822; Q9HMQ3;
DC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-2 chain precursor.
GN CHRNA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=97062879; PubMed=8906617;
RA Elliott K.J., Ellis S.B., Bercshan K.J., Urrutia A.,
RA Chavez-Noriega L.E., Johnson E.C., Velicelceti G., Harpold M.M.;
RT "Comparative structure of human neuronal alpha 2, alpha 7 and beta
RT 2-beta 4 nicotinic acetylcholine receptor subunits and functional
RT beta 4 subunits.";
RT J. Mol. Neurosci. 7:217-228(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Groot Kormelink P.J.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Blechschmidt K., Rosenthal A.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC MEMBRANE.
CC -1- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT TYPES
CC OF SUBUNITS: ALPHA AND NON-ALPHA (BETA). ALPHA-2 SUBUNIT CAN BE
CC COMBINED TO BETA-2 OR BETA-4 TO GIVE RISE TO FUNCTIONAL RECEPTORS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U62431; AAB40109.1; -
DR EMBL; Y16281; CAA76154.1; -
DR EMBL; AF311103; -, NOT_ANNOTATED_CDS.

```

```

DR MIM; 118502; -
DR InterPro: IPR001188; GABA_receptor.
DR InterPro: IPR001175; Neur_channel.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_membr; 1.
DR PRINTS: PR00252; NRIONCHANNEL.
DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
DR Postsynaptic membrane; Ionic channel; Glycoprotein; signal;
KW Transmembrane; Multigene family.
FT SIGNAL 1
FT CHAIN 26
FT DOMAIN 27 529
FT TRANSMEM 27 264
FT TRANSMEM 265 289
FT TRANSMEM 297 315
FT TRANSMEM 331 352
FT DOMAIN 353 502
FT TRANSMEM 503 521
FT TRANSMEM 503 521
FT DISULFID 183 248
FT DISULFID 247 248
FT CARBOHYD 79 79
FT CARBOHYD 129 129
FT CARBOHYD 235 235
FT CONFLICT 125 125
FT CONFLICT 125 125
SQ SEQUENCE 529 AA; 59735 MW; 7F512B06CDD9A9FD CRC64;

Query Match 18.0%; Score 416; DB 1; Length 529;
Best Local Similarity 27.6%; Pred. No. 6e-26;
Matches 109; Conservative 75; Mismatches 153; Indels 58; Gaps 12;

QY 13 GPCPLLLQML-MGMLMLQTSVPGATATADPKNA-----NKAIDR 53
Db 2 GPCPVFLSFTKLSLMMWLLTPAGGEAKRPPRAGDPLSSPSPTALPQGGSHETEDR 61
QY 54 LHAQLETNVDSVDPVFGCTPTNVSLSENVYV-----IDDELNGKLTTCWNLNMR 106
Db 62 LKHLFLRGYNRRARVP-----PNTSDVYVIRGSLSTIAQLIDYDEKQMMTINWVWQEHNS 116
QY 107 DEERVWQPSQYDNTIQTLKSEVWTPQTLFNGDEGGGLMAETQVYLSH--DGHRMP 163
Db 117 DYKLRWNPADFCNITSLRVSEMIWIPDIVLNNDGE-FAVTHMTKAHLFTGTVHWVP 175
QY 164 PAVYTAFCCLNMLNMPHDQSKLKISW---GLKVLVPENGTARGESLDHDDLVSQPEW 220
Db 176 PAIYKSSCIDYVTFPPDOONCKMKFGSWTYDKAKIDLQ---MQQIVDLKDYESGEM 231
QY 221 EIVDSRAHFVSODY-----YGYMEYTLFAORSSMYTAVIYTPASCIIVLALSAFWLP 274
Db 232 AIVNATGTYNSKKYDCCALITPDVYAFVIRLPLFYTIMLIPCLLISCLTVLVEFLPS 291
QY 275 HMGGEKIMINGLLITVIAFLMFAQLLPVLSNNTPLVVFYTSLSLYSVSTIVEVLV 334
Db 292 DC-GEKITLICISVLSTLVEFLLETIELPSTSLVPLIGELYLFTMLFVLSIVITVEVL 350
QY 335 YLATGKHKR-----RLPEALRKLLHCHLCTWLLIS 364
Db 351 NV-----HHRSPSTHMPHWVHGALLGVPRWLLMN 381

RESULT 6
ACHA_RAT
ID ACHA_RAT STANDARD: PRT: 457 AA.
AC P25108;
DC 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE Acetylcholine receptor protein, alpha chain precursor.
GN ACHR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RX MEDLINE=91099317; PubMed=1702709;  
 RA Witzmann V., Steh E., Barg B., Konno T., Koenen M., Kues W.,  
 RA Chlado M., Hofmann M., Sakmann B.;  
 RT "Primary structure and functional expression of the alpha-, beta-,  
 RT gamma-, delta- and epsilon-subunits of the acetylcholine receptor  
 RT from rat muscle.";  
 RL Eur. J. Biochem. 194;437-448(1990).  
 CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN  
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND  
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA  
 CC MEMBRANE.  
 CC -1- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, AND ONE EACH OF THE BETA,  
 CC DELTA, AND GAMMA (IN IMMATURE MUSCLE) OR EPSILON (IN MATURE  
 CC MUSCLE) CHAINS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X74832; CA52826.1; -  
 DR PIR: S13872; S13872.  
 DR InterPro: IPR000188; GABAA\_receptor.  
 DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
 DR Pfam: PF02932; Neur\_chan\_mem; 1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; 1.  
 DR Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
 KW Transmembrane.  
 KM  
 FT SIGNAL 1 20  
 FT CHAIN 21 457  
 FT DOMAIN 21 230  
 FT TRANSMEM 231 235  
 FT TRANSMEM 263 281  
 FT TRANSMEM 297 316  
 FT DOMAIN 317 428  
 FT TRANSMEM 429 447  
 FT DISULFID 148 162  
 FT DISULFID 212 213  
 FT CARBOHYD 161 161  
 FT SEQUENCE 457 AA; 51866 MW; 776AE3B8DF8F683 CRC64;  
 SO  
 Query Match 17.98; Score 414.5; DB 1; Length 457;  
 Best Local Similarity 25.4%; Pred. NO. 6.6e-26;  
 Matches 118; Conservative 99; Mismatches 174; Indels 73; Gaps 19;  
 OY 20 LQMLGMLGIVTVPATADPKNANVKALDLHGLFNNYSDQPV---FQGIPTN 76  
 DB 1 MELTAVILLGLGCSA-STVIGASEHET-----RLVAKLEKDYSSVVRPGDREIYQVT 52  
 OY 77 VLEEMVVTYIDIDELNGKLTTHCWLNRWDEERWOPQVDNTQTILKSEVWTPQIT 136  
 DB 53 VGLQL-IQLINVDENVQITTVNRLKQOWDYNLKAMPDQYGVKAKHIPEKIKRDEV 111  
 OY 137 LFNQDEG--GLAETQVTLSDHGFRRMPAPVATYACELANLNPNHDKQSCIKLIGSW-- 192  
 DB 112 LYNNAEDFAIVKFTKLLDYTGHTWTPPAIFKSYCEIIVTHPPDEQNSMKLGTWY 171  
 OY 193 --GLKAVLPENGTAEGSLDHDLDVQSPWEIVDSRA--HFVSDYYG-----YME--Y 240

DB 172 DGSVAINPES-----DQPLDSNFMESGENVKEARKWKHNV---FYSCCPNTPYDITY 223  
 OY 241 TITPAQRSSMTAVIYTPASCTVILASAPFLPMPHMGKIMINGLLIYIAAFMYFAQ 300  
 DB 224 HFVQGRPLIVFIYVILICPLFSFLTSLVFLYLPD-SGKMTLSISLSTVFLVIVE 282  
 OY 301 LLLPVSNTPLVVIFSTSLIYSVSTIVEVLYLATGKHRR-----LPEALRKLHG 355  
 DB 283 LIPSTSAVPLICKYMLFTWVFYIASIITIVIVY----NTHHRSPTTHMPEWVRKVID 338  
 OY 356 HLGWLLSVFSTGESAQK-----TKEMDHPREADEQSSPLG-----INHEVGA- 406  
 DB 339 TTPIMFSTYMKRPSRDKRIFTEIDIS--DISKPGPPPGHSPLIKPEVSAI 396  
 OY 407 -----KAND-----FDNALLATAVDRISFVSFAFLI 434  
 DB 397 EGVYIAETMKSDQSNNSASEMKYVAMVDHLLGVFMLVCL 440  
 RESULT 7  
 ACH6\_RAT STANDARD; PRT; 493 AA.  
 AC PA3143;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Neuronal acetylcholine receptor protein, alpha-6 chain precursor.  
 GN CHRN6 OR ACR6.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Boulter J.;  
 RU Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN  
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND  
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA  
 CC MEMBRANE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L08227; AAA41674.1; -  
 DR HSSP: P01270; 1HPH.  
 DR InterPro: IPR000188; GABAA\_receptor.  
 DR InterPro: IPR001175; Neur\_channel.  
 DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
 DR Pfam: PF02932; Neur\_chan\_mem; 1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; 1.  
 KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
 KM Transmembrane; Multigene family.  
 FT SIGNAL 1 30  
 FT CHAIN 31 493  
 FT DOMAIN 31 240  
 FT TRANSMEM 241 265  
 FT TRANSMEM 272 290  
 FT TRANSMEM 306 327  
 FT DOMAIN 328 464  
 FT TRANSMEM 465 484  
 FT DISULFID 158 172  
 FT DISULFID 222 223  
 FT  
 BY SIMILARITY.  
 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,  
 ALPHA-6 CHAIN.  
 EXTRACELLULAR (POTENTIAL).  
 POTENTIAL.  
 POTENTIAL.  
 CYTOPLASMIC (POTENTIAL).  
 POTENTIAL.  
 BY SIMILARITY.  
 ASSOCIATED WITH RECEPTOR ACTIVATION  
 (BY SIMILARITY).



Db 289 SAVFLIKYMLFTWVFVIAIITIVYI---NTHHSPSTHINPEWKRKVEIDTIPNIM 344

QY 362 LLSVSTTGSGQAEK---TKEMDEHPYEADDEQSSPLG-----INHTVEPGA----- 406

Db 345 FFSFMKRPNRKQKRIETFEIDIS--DISGKPGPPMGPSPILKHPVMSALEGVYKI 402

QY 407 ----KANO-----FDWALLATAVDRISFVFSFSLATLI 434

Db 403 AETWKSQDESNNAEWEKRYVAMVDHILLGVMLVCLII 440

RESULT 9

ACHO\_CHICK ID ACHO\_CHICK STANDARD; PRT; 455 AA.

AC P43679;

DT 01-NOV-1995 (Rel. 32, Created)

DT 16-OCT-2001 (Rel. 40, last sequence update)

DT 16-OCT-2001 (Rel. 40, last annotation update)

DE Neuronal acetylcholine receptor protein, beta-3 chain precursor.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

OC NCBI\_TaxID=9031;

OX [1]

RN SEQUENCE FROM N.A.

RP STRAIN=WHITE LEGHORN;

RX MEDLINE=95155414; PubMed=7852408;

RA Hernandez M.C., Erkman L., Matter-Sadzinski L., Roztocil T., Ballivet M., Matter J.M., "Characterization of the nicotinic acetylcholine receptor beta 3 gene. Its regulation within the avian nervous system is effected by a promoter 143 base pairs in length.";

RT J. Biol. Chem. 270:3224-3233(1995).

RL [2]

RN REVISION TO 331.

RA Ballivet M.;

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA MEMBRANE.

CC -1- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT TYPE OF SUBUNITS: ALPHA AND BETA.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: RELATIVELY ABUNDANT IN THE DEVELOPING RETINA AND IN THE TRIGEMINAL GANGLION.

CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; X83739; CAB59814.1; -

DR InterPro: IPR000188; GABAA\_receptor.

DR InterPro: IPR000175; Neur\_channel.

DR Pfam: PF02931; Neur\_chan\_Lbd; 1.

DR Pfam: PF02932; Neur\_chan\_memb; 1.

DR PRINTS: PR00252; NRIONCHANNEL.

DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; 1.

KM postsynaptic membrane; Ionic channel; Glycoprotein; Signal;

KW Transmembrane; Multigene family

FT SIGNAL 1 20

FT CHAIN 21 455

FT DOMAIN 21 229

FT TRANSMEM 230 254

FT TRANSMEM 262 279

FT TRANSMEM 296 317

FT DOMAIN 318 425

FT TRANSMEM 426 444

FT DISULFID 150 164

FT CARBOHYD 48 48

FT CARBOHYD 163 163

SO SEQUENCE 455 AA; 52276 MW; 3D7A7ARA7D8511A CRC64;

Query Match 17.7%; Score 410; DB 1; Length 455;

Best local Similarity 26.3%; Pred. No. 1.5e-25;

Matches 115; Conservative 85; Mismatches 167; Indels 70; Gaps 17;

QY 52 DRCHAGLFTWVDSQVDFVFGTPT--NVSLEMYVT--YIDDELNGKLTTHQWMLRWDEE 109

Db 27 DALRLHFLFOGQYQKWRREVNSNDTIKVLFLKISQLDVDEKQMLMTNWMLAQEWMDHK 86

QY 110 RVMPQSDYDNTQTTLTKSSVWTPQITLF--NGD---EGGIMAEFQVTLSDGHFRMRPPA 165

Db 87 LSMNPEEYGGITAIRVSESLMLPDIYLFENADGREGSLM--TKAIVKNGVQWMPA 144

QY 166 VYTAVCELMNLMNPHDKQCKLIGSWG-----LKVLPENGTARGESLDHDLVQSPW 220

Db 145 SYKSSCTMELTFPPFDQNSMKFGSWTYDGSWMDLVD-----ENVDTKPFQNGEM 198

QY 221 EYDSRAHEVSQD---YGYMEYTLTAQRSSMYTAVIYTPASCIYIALSAFWLPPHM 276

Db 199 EILNAKMGKGNRKDGILSYFPFYTSFVLRRLPLEYTLFLIPCLIGLSFLTVLVEYLPSPD- 257

QY 277 GGERIMINGLLIYIAFLWYFNQDLVLSNNPLVYIFSTLLYTSVSTIYEVLYL 336

Db 258 EGKLSLSTSVLSVLVLLVLEIEIIPSSKVPILGEVLYLFIETLIIYTVFIVNV 317

QY 337 ATGKHKR-----RLPEALRKLLHGLTWLLSVSTTGSGQAE-KT 377

Db 318 ----HHRSSATYHPMARWYKRLFLQKLPRL--CMKHVDVRY----SFSDTEKETTLKS 367

QY 378 KEMDEHPYEADDEQSSPLGINHTVEPG-----AKANP-----DWALLATAVDRIS 424

Db 368 LKPKQKHKQAKQDEKVFIAFLERKADSIYRSRHKKDAFIRQVQWDMKFAVQVLDRI- 426

QY 425 FVFSFLAFLIATRCVS 441

Db 427 ----FLWLELVAVSYTGSV 440

RESULT 10

ACHX\_HUMAN ID ACHX\_HUMAN STANDARD; PRT; 450 AA.

AC Q9GZ26;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, last sequence update)

DT 01-MAR-2002 (Rel. 41, last annotation update)

DE Neuronal acetylcholine receptor protein, alpha-10 chain precursor (nicotinic acetylcholine receptor subunit alpha 10) (NACHR alpha 10).

GN CHRNA10 OR NACHRA10.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OC NCBI\_TaxID=9606;

OX [1]

RN SEQUENCE FROM N.A.

RP TISSUE=Skeletal muscle;

RA Sgar F., Charpentier E., Bertrand S., Walker N., Agnel M., Graham D., Caput D., Bertrand D., Besnard F.;

RT "Cloning and expression of a novel nicotinic alpha10 subunit from human that confers functionality to the alpha9 subunit.";

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Inner ear, and Tonsil;

RX MEDLINE=21248672; PubMed=11350119;

RA Lustig L.R., Peng H., Hiel H., Yamamoto T., Fuchs P.A.;

RT "Molecular cloning and mapping of the human nicotinic acetylcholine

DB	Query Match	Best Local Similarity	Match 127: Conservative	Score 409: DB 1: Length 450: Pred. No. 1.8e-25; Mismatches 188; Indels 62; Gaps 18;
DB	123 DAOPGSASTNVNLRHDAVRWDADAIRRSSCRDVVAAPFPAQHCGLTFGSWTHGHQL 182	17.7%	78	
CC	receptor alpha10 (CHRNA10)."			
RL	Genomics 73:272-283(2001).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Vandenberk I., van der Helm L., Nieuwstraten D.M., Verfaille C.,			
RA	Kremer A., van der Spek P., Masure S., Hoefnagel E., Yon J.,			
RA	Groot Kormelink P.J., Luyten W.H., Grantham C.J.:			
RT	"Identification and functional characterisation of a novel human			
RT	neuronal nicotinic acetylcholine receptor subunit alpha 10.";			
RL	submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.			
CC	-1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN			
CC	EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND			
CC	LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA			
CC	MEMBRANE.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- TISSUE SPECIFICITY: Expressed in inner-ear tissue, tonsil,			
CC	immortalized B-cells, cultured T-cells and peripheral blood			
CC	lymphocytes.			
CC	-1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ).			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL: AJ278118; CAC20435.1; -			
DR	EMBL: AF199235; AAC00795.2; -			
DR	EMBL: AF327367; AAK14333.1; -			
DR	EMBL: AJ295237; CAC16144.1; -			
DR	MM: 606372; -			
DR	InterPro: IPR000188; GABA_A_receptor.			
DR	InterPro: IPR001175; Neur_channel.			
DR	Pfam: PF02931; Neur_chan_LBD, 1.			
DR	Pfam: PF02932; Neur_chan_membr, 1.			
DR	PRINTS: PR00252; NRIONCHANNEL.			
DR	PROSITE: PS00236; NEUROTR_ION_CHANNEL, 1.			
KW	Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;			
KW	Transmembrane; Multigene family.			
FT	SIGNAL	1	24	POTENTIAL.
FT	CHAIN	25	450	
FT				POTENTIAL.
FT	DOMAIN	25	237	NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
FT	TRANSMEM	238	258	ALPHA-10 CHAIN.
FT	TRANSMEM	268	288	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	302	322	POTENTIAL.
FT	DOMAIN	323	428	POTENTIAL.
FT	TRANSMEM	429	449	CYTOPLASMIC (POTENTIAL).
FT	DISULFID	154	168	POTENTIAL.
FT	DISULFID	218	219	POTENTIAL.
FT				BY SIMILARITY.
FT				ASSOCIATED WITH RECEPTOR ACTIVATION
FT				(BY SIMILARITY).
FT	CARBOHYD	40	40	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	56	56	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	450 AA;	49704 MM;	0A1DAD32D9069870 CRC64;

Qy	197	-VLPENGTARSEDJDDHLOSPEKEIYDSRA-----HFVSQDYGVMEYTLTAQRSS	249
Dy	183	DVRP-----RGAASLADFEVNEVRVGMPPRRRRVLTGGCSEPPDVTTLRRNA	237
Qy	250	MYTAIVTPASCIYVLALSAFWLPNPMGKERIMINGLIIVIAAFWMEFAOLLPLVSNNT	309
Dy	238	AYVCLLLLPVLLISLAPLAFHLPAD-SGEKVSGLVTLVLLATYFQOLLASMP-ASV	295
Qy	310	PLVVFVSTSLIYLSVSTIVEVLYLT-ATGKHKRRLDEALRKLLHGLTWLLISVST	368
Dy	296	PLIGKYVATATMTWTFSTALTITLIMNLHYCGPSVPYPMARALLHGLHLAGLCV---RE	352
Qy	369	TGESQAE-KTFEMDHPPEADEDEQSSPLG-----INHFVGAKANOP--	411
Dy	353	RGEPCGGSRPPELSP-QSPEGAGAPPAGCHEPRCLCRQDALDHY---ATTANTFRS	408
Qy	412	-----DWALLATAVDRISEVFSLSAPLILAI	437
Dy	409	HRAQRCHEDWKRLARVMDREFLAIFEFMALVMSL	443
RESULT	11		
ACH4-CHICK			
ID	ACH4.CHICK	STANDARD;	PRT; 622 AA.
AC	P09482:		
DT	01-MAR-1989	(Rel. 10, Created)	
DT	01-MAR-1989	(Rel. 10, Last sequence update)	
DT	16-OCT-2001	(Rel. 40, Last annotation update)	
DE	Neuronal acetylcholine receptor protein, alpha-4 chain precursor.		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OC	Gallus.		
OX	NCBI_TaxID=9031;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RX	MEDLINE=88283624; PubMed=3267226;		
RA	Nef P., Oneysen C., Alliod C., Couturier S., Ballivet M.;		
RT	"Genes expressed in the brain define three distinct neuronal		
RT	nicotinic acetylcholine receptors.";		
RL	EMBO J. 7:595-601(1988).		
RN	[2]		
RP	MUTAGENESIS OF GLU-289, AND SUBUNITS.		
RX	MEDLINE=91172320; PubMed=2005979;		
RA	Cooper E., Couturier S., Ballivet M.;		
RT	"Pentameric structure and subunit stoichiometry of a neuronal		
RT	nicotinic acetylcholine receptor.";		
RL	Nature 350:235-238(1991).		
CC	-I- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN		
CC	EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND		
CC	LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA		
CC	MEMBRANE.		
CC	-I- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT		
CC	TYPE OF SUBUNITS: ALPHA AND NON-ALPHA (ALSO CALLED BETA). A		
CC	FUNCTIONAL RECEPTOR SEEMS TO CONSIST OF TWO ALPHA-CHAINS AND		
CC	THREE NON-ALPHA CHAINS.		
CC	-I- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-I- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/annouce/ib-sib.ch">http://www.isb-sib.ch/annouce/ib-sib.ch</a> ).		
CC	or send an email to <a href="mailto:license@ib-sib.ch">license@ib-sib.ch</a> ).		
CC	-----		
DR	EMBL; X07348; CAA30285.1; -;		
DR	EMBL; X07349; CAA30285.1; JOINED.		
DR	EMBL; X07350; CAA30285.1; JOINED.		
DR	EMBL; X07351; CAA30285.1; JOINED.		

DR EMBL: X07352; CAA30285.1; JOINED.  
DR EMBL: X07359; CAA30285.1; JOINED.  
DR EMBL: AJ250361; CAB59626.1; -.  
DR PIR: S00379; ACCH4N.  
DR InterPro: IPR000188; GABAA\_receptor.  
DR InterPro: IPR001175; Neur\_channel.  
DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
DR Pfam: PF02932; Neur\_chan\_memb; 1.  
DR PRINTS: PR00252; NRIONCHANNEL.  
DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; 1.  
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
KW Transmembrane; Multigene family.  
FT SIGNAL 1 23  
FT CHAIN 24 622 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,  
FT DOMAIN 24 262 ALPHA-4 CHAIN.  
FT TRANSMEM 238 262 EXTRACELLULAR.  
FT TRANSMEM 270 288  
FT TRANSMEM 304 325  
FT DOMAIN 326 359 CYTOPLASMIC.  
FT TRANSMEM 596 614  
FT DISULFID 156 170  
FT DISULFID 220 221  
FT CARBOHYD 52 52 BY SIMILARITY.  
FT CARBOHYD 102 102 ASSOCIATED WITH RECEPTOR ACTIVATION  
FT CARBOHYD 289 289 (BY SIMILARITY).  
FT MOTABRN E-K: REDUCES CHANNEL CONDUCTANCE BY  
FT HALF.  
SQ SEQUENCE 622 AA; 70837 MW; 290B035893D6485A CRC64;

Query Match 17.6%; Score 407.5; DB 1; Length 622;  
Best Local Similarity 29.4%; Pred. No. 3.6e-25;  
Matches 96; Conservative 73; Mismatches 119; Indels 39; Gaps 10;  
QY 50 ALDLHAGLFTNYSDQVPGTPTNYSLEMYTY-----IDIDELNKLTHCHLN 102  
DB 31 AEEPLKLLKSGYKMSRPV-----ANISDVLPVFGLSIAQLIDVDKKNMTTNVWK 85  
QY 103 LRMDEERWVQPSQYDNTQTLLKSEVWTPQITLFNDEGLMAETQVTSI--DGHF 159  
DB 86 QEMHDYLRMPQREYEVNTSIRISBELMRDYLNNADD-FAVHILTHAHLFYDORI 144  
QY 160 RMPFPAVYTAACELNMLNMPHDKOSCKIKGSW---GLKVVLPENGARGESLDHDLVQ 216  
DB 145 KMPFPAVYKSSCIDVFEFPDQONCKMKFSGWTYDKAKIDL---VSMHSHVDQLDYWE 200  
QY 217 SPEPEIYDSRAHEVSDY-----YGYMEYTLTAQRSSMYTAVIYTPASCIIVILASAF 270  
DB 201 SGEVVIINAGVNSKKECTEITPDTYSFTIRRLPELFTINLILPCLLSCTLVLF 260  
QY 271 WLPFHMGEKIMINGLLIIVIAFLMYFAQLLPVLSNNTPLVVFYSTLLLSVSTIVE 330  
DB 261 YLPSEC-GEKTLICISVLSTFLVLLITELIIPSTSLVPLIGRYLFTMLFVLSITIT 319  
QY 331 VLVLYLATGKRK-----RLPEALRKL 352  
DB 320 VEVLVN-----HHRSPRTHTPMDWVRV 342

RESULT 12  
ACHI\_SCHGR STANDARD; PRT; 557 AA.  
AC P23414;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE Acetylcholine receptor protein, alpha-L1 chain precursor.  
OS Schistocerca gregaria (Desert locust).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Plekoyota; Neoptera; Orthopteroidea; Orthoptera; Caellifera;  
OC Acridomorpha; Acridoidea; Acrididae; Schistocerca.  
OX NCBI\_Taxid=7010;

RNA SEQUENCE FROM N.A.  
RX MEDLINE=91092263; PubMed=1702381;  
RA Marshall J., Buckingham S.D., Shingai R., Lunt G.G., Goosey M.W.,  
RA Darlison M.G., Sattelle D.B., Barnard E.A.;  
RT "Sequence and functional expression of a single alpha subunit of an  
RT insect nicotinic acetylcholine receptor.";  
RL Embo J. 9:4391-4398(1990).  
CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN  
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND  
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA  
CC MEMBRANE.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X55439; CAA39081.1; -.  
DR PIR: S12359; S12359.  
DR InterPro: IPR000188; GABAA\_receptor.  
DR InterPro: IPR001175; Neur\_channel.  
DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
DR Pfam: PF02932; Neur\_chan\_memb; 1.  
DR PRINTS: PR00252; NRIONCHANNEL.  
DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; 1.  
DR Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
KW Transmembrane; Multigene family.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 557 ACETYLCHOLINE RECEPTOR PROTEIN, ALPHA-L1  
FT DOMAIN 24 244 CHAIN.  
FT TRANSMEM 245 266 EXTRACELLULAR.  
FT TRANSMEM 274 294  
FT TRANSMEM 308 329  
FT DOMAIN 330 500  
FT TRANSMEM 501 523 CYTOPLASMIC.  
FT DISULFID 151 165  
FT DISULFID 224 225  
FT CARBOHYD 47 47 BY SIMILARITY.  
FT CARBOHYD 235 235 ASSOCIATED WITH RECEPTOR ACTIVATION  
FT DOMAIN 382 400 (BY SIMILARITY).  
FT DOMAIN 406 422 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT SIGNAL 422 422 ALA/SER-RICH.  
SQ SEQUENCE 557 AA; 63026 MW; 168389C887DFDF3E CRC64;

Query Match 17.6%; Score 407; DB 1; Length 557;  
Best Local Similarity 23.3%; Pred. No. 3.4e-25;  
Matches 128; Conservative 91; Mismatches 176; Indels 154; Gaps 19;  
QY 17 PLLDMLMGLMGLTSVPGATATADPKMANVKAIDRLHAGLFTNYSDQVPGTPTN 76  
DB 7 PMLL-----LLLLLLHNR-AAANPDAK-----RLYDLSNRYRLLRPVSNNTDT- 51  
QY 77 VSELMVVTYIDIDELNKG---LTHCHWLNLRMDEERWQPSQYDNTQTLLKSEVWTP 133  
DB 52 VLVKGLRLSQLIDLNKDOIITTNVWLEHWDHKKRMPDAEYGVTELYVPEHIMLP 111  
QY 134 QITLFNDEGLMAE--TQVTLSDHGFRMPRPVYTAACELNMLNMPHDKOSCKIKGS 191  
DB 112 DIVLYNNADEGYVTYTKAVLHHTGKVWTPPAIFKSCSIDVRYFPDQOTCEMKFGS 171  
QY 192 WGLKVVLPENGARGESLD-----HDD-----LVOSPMEIYDSRAHEVSOD 233  
DB 172 W-----TYGGDQDLKHNQKDDKKVKKQIGIDLREYVEMDILIGVAREKREKY 221  
QY 234 Y-----YGYMEYTLTAQRSSMYTAVIYTPASCIIVILASAFWLPPHMGGEKIMINGLL 287

```

DB 222 YPCSAEPDIPDIFNTLRKTLFETVNLIVPCVGISYLVFLYPAD-SGEEKALCISI 280
OY 288 IIVIAAFMYFAFOLLPLVSNNTPLVVFSTSLVLSVSTIVEVL---YLATGKHK-- 342
DB 261 LISQTMFLLSLSEITPSTSLALPLGKYLFTWVLVGLSVTTITVLNVYKRPSTHKMA 340
OY 343 -----RRLPEAL-----RKLLHGH---LGTWLLSVSTYSGESQAEK 376
DB 341 PWRKVFIRRLPKLLMLRVPEQLDLASKRLLRHAHNSKLSAAAAAAYAAASSAASS 400
OY 377 TKEMDEHPREEADEQESSFLGINTHEVPG-----AKANQF----- 411
DB 401 PDSLKHHLHQ--HGHNLQLHNLQRPQCGNGLHSATNRFEGSGAGFGLPSVVGLDGS 458
OY 412 -----DMALLNPAVDRISFVSFSLAF 432
DB 459 LSDVATRKYPPELEKALHNVLFIQNHMORODEPAEDQDMGVAMVLDRLFLMITIAS 518
OY 433 LI--LAIRC 439
DB 519 IYGTFAIIC 527

RESULT 13
ACH4_RAT STANDARD: PRT: 630 AA.
AC P09483: 035769;
DT 01-MAR-1989 (Rel. 10, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-4 chain precursor.
GN CHRM4 OR ACRA4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Hypothalamus, and Hippocampus;
RX MEDLINE=87159533; PubMed=3829125;
RA Goldman D.J., Denneris E.S., Luyten W., Kochhar A., Patrick J.,
RA Heinemann S.F.;
RT "Members of a nicotinic acetylcholine receptor gene family are
RT expressed in different regions of the mammalian central nervous
RT system."
RL Cell 46:965-973(1987).
RN [2]
RP REVISIONS.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RA Hatley M., Goldman D.J., Heinemann S.F.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Boulter J., Denneris E.S., Evans K., Heinemann S.F.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP PARTIAL SEQUENCE OF 31-47.
RX MEDLINE=87276531; PubMed=3609304;
RA Whiting P., Esch F., Shimazaki S., Lindstrom J.;
RT "Neuronal nicotinic acetylcholine receptor beta-subunit is coded for
RT by the cdna clone alpha 4."
RL FEBS Lett. 219:459-463(1987).
CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC MEMBRANE.
CC -1- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT TYPES
CC OF SUBUNITS: ALPHA AND NON-ALPHA (BETA). ALPHA-4 SUBUNIT CAN BE
CC COMBINED TO BETA-2 OR BETA-4 TO GIVE RISE TO FUNCTIONAL RECEPTORS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA-4-1 (SHOWN HERE) AND
CC ALPHA-4-2; ARE PRODUCED BY ALTERNATIVE SPLICING.

```

```

CC -1- TISSUE SPECIFICITY: IN VARIOUS REGIONS OF THE CENTRAL NERVOUS
CC SYSTEM.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC DR EMBL: M15682; AAA41676.1; -.
CC DR EMBL: L31620; AAC97071.1; -.
CC DR EMBL: AF007212; AAB64439.1; -.
CC DR PIR: A26456; A26456.
CC DR InterPro: IPR000188; GABAA_receptor.
CC DR InterPro: IPR001175; Neur_chan.
CC DR Pfam: PF02931; Neur_chan_LBD; 1.
CC DR Pfam: PF02932; Neur_chan_mem; 1.
CC DR PRINTS: PR00252; NRIONCHANNEL.
CC DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
CC DR PostSynaptic membrane; Ionic channel; Glycoprotein; Signal;
CC KM Transmembrane; Multigene family; Alternative splicing.
CC FT SIGNAL 1 30
CC FT CHAIN 31 630
CC FT DOMAIN 32 249
CC FT TRANSMEM 250 270
CC FT TRANSMEM 279 299
CC FT TRANSMEM 313 333
CC FT DOMAIN 334 604
CC FT TRANSMEM 605 625
CC FT DISULFD 163 177
CC FT DISULFD 227 228
CC FT CARBOHYD 59 59
CC FT CARBOHYD 109 109
CC FT CARBOHYD 176 176
CC FT VARSPPLIC 629 630
CC FT CONFLICT 1 4
CC FT CONFLICT 58 58
CC FT CONFLICT 196 196
CC FT CONFLICT 301 301
CC FT SEQUENCE 630 AA; 70192 MW; 1DFC9E1378E5417 CRC64;

Query Match 17.6%; Score 407; DB 1; Length 630;
Best Local Similarity 25.9%; Pred. No. 4e-25;
Matches 120; Conservative 84; Mismatches 176; Indels 84; Gaps 14;

OY 12 SGPGLPLLLQMLGMLGLTSPGATATADKNNVKAHLHAGLFTNYDSQVQPYQ 71
DB 4 SGTGAPRPLLLPLLLLTGTLPPASSHETFRANHEERLLKR---LFGYKMSRPV-- 57
OY 72 GRTNVSLSEMYTY-----IDIDELNKLTTCHLNLNRMDERVQWQSOYDINTQTT 124
DB 58 ---ANISDVVLVFGSLAQLLIDVDEKQMTTNVWVKQEMDYKLRMDPGDYENVTSIR 114
OY 125 LKSESEVTPQITLNFNGDEGGLAEQVTLSH---DGHFWMRPAYVTAACEMLNMPHD 181
DB 115 ISELIMRDIYLVYNNADD-FAYVHLTAHLFYDGRQVMTPTAIYKSSCIDYFFPPD 173
OY 182 KOSCKIKIGSW--GLKVLVPENGARGESLDHDLVQSPWEIYDSRAHFVSODY---- 234
DB 174 QONCTPKFGSMTYDRAKIDL---VSMHSRYDQLDFWEGSEWIVDAVGYWTRKYECA 229
OY 235 --YGMETVTLNQRSSMTAVIYTPASCIYVLLASAFLLPRHMGKERKINGLLIYIA 292
DB 230 EIVPDITVAFTIRRLPLEFTYIMLIPCLLISCLIVLVFLYSEC-GEKYTLCISVLSLT 288
OY 293 AFLWYFAQLLP-VLSNNTPLVVFSTSLVSVSTIVEVLVLYLATGKHK----- 342

```

```

Db 289 VFLILITELIPSTSLVPLIGELFTMLFVTLISIVTVLVNW---HHRSPTHTMPA 345
QY 343 -----RLPEALKRLHG-----HLGTWLLSVFSTTG 370
Db 346 WVRVFDIDYPRLLFMKRPSPVKNCRLLIESMKMANAPFWEPPGEGPILSDICNOG 405
QY 371 ESQAETKREMDHPYEADEQSSPLGINHTVPGANOPDMA 414
Db 406 LSPAPFCNPFDVAVETQPTCRSPPL-----EVPDLKTSVEKA 444

RESULT 14
5HT3_HUMAN STANDARD; PRT; 478 AA.
ID 5HT3_HUMAN
AC P46098;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE 5-hydroxytryptamine 3 receptor precursor (5-HT-3) (serotonin-gated ion
channel receptor) (5-HT3R).
GN HTR3A OR HTR3 OR 5HT3R.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RX MEDLINE=96018832; Pubmed=7565620;
RA Miyake A., Mochizuki S., Takemoto Y., Akuzawa S.;
RT "Molecular cloning of human 5-hydroxytryptamine3 receptor:
heterogeneity in distribution and function among species.";
RT Mol. Pharmacol. 48:407-416(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Amalgala;
RX MEDLINE=99127223; Pubmed=9928262;
RA Bruns M., Gothert M., Hayer M., Bonisch H.;
RT "Molecular cloning of alternatively spliced human 5HT3 receptor
cDNAs.";
RT Ann. N.Y. Acad. Sci. 861:234-235(1998)
CC -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR
5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTION
AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THIS RECEPTOR IS
A LIGAND-GATED ION CHANNEL, WHICH WHEN ACTIVATED CAUSES FAST,
DEPOLARIZING RESPONSES IN NEURONS. IT IS A CATION-SPECIFIC, BUT
OTHERWISE RELATIVELY NONSELECTIVE, ION CHANNEL.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: DA9394; BAA08387.1; -
CC EMBL: AJ003079; CA05852.1; -
CC MIM: 182139; -
CC InterPro: IPR000188; GABAA_receptor.
CC InterPro: IPR001175; Neur_channel.
CC Pfam: PF02931; Neur_chan_LBD.1.
CC Pfam: PF02932; Neur_chan_membr.1.
CC PRINTS: PR00252; NRIONCHANNEL.
CC PROSITE: PS00236; NEUROTR_ION_CHANNEL.1.
CC Ionic channel; Transmembrane; Receptor; Glycoprotein; Signal;
KM Phosphorylation.
FT SIGNAL 1 23
FT CHAIN 24 478 POTENTIAL.
FT DOMAIN 24 241 5-HYDROXYTRYPTAMINE 3 RECEPTOR.
EXTRACELLULAR (POTENTIAL).

```

```

FT TRANSMEM 242 268 1 (POTENTIAL).
FT DOMAIN 269 273 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 274 292 2 (POTENTIAL).
FT DOMAIN 293 302 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 303 321 3 (POTENTIAL).
FT DOMAIN 322 455 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 456 475 4 (POTENTIAL).
FT DOMAIN 476 478 EXTRACELLULAR (POTENTIAL).
FT DISULFID 157 171 BY SIMILARITY.
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 170 170 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 478 AA; 55280 MW; 24CA9A232286FBC9 CRC64;

Query Match 17.5%; Score 404.5; DB 1; Length 478;
Best Local Similarity 26.5%; Pred. No. 4.5e-25;
Matches 126; Conservative 82; Mismatches 194; Indels 73; Gaps 18;

QY 21 QMLGMILMGITSVGATATADPKNAVVKALDRHAGLFTNYDSOVQPVQ-GPTNVSL 79
Db 7 QALLALLPPTLLAAGEARRS---RNTTRPALRLSDYLLNRYKGVAPVRDMKRPITVSI 63
QY 80 EMVY-TYIDIDELNGKLTTHQWMLRNRDERVWQSQYDNITQITLKSEVWTPTTLF 138
Db 64 DVIYVAILNVDENKQVLTITWYRQVWTDQLQNNPDPDMITKLSIPTDSINVPDI-LI 122
QY 139 NG--DEGLAAETQVYTLSDHGHFRMPAYTAVCELMNMPDKOSCKIKISGSLKV 196
Db 123 NEFVDVKSPPNIPVYVYRHOGEVQNYKPLQVYVTAQSDIDYFPPDVQNCSTLFTSW-LHT 181
QY 197 V-----LPNGTARGESLDHDDLVQSPPEIVYSRAHF-----VSODYGYMEYTL 242
Db 182 IQDINISLMRLPEKYS-----DRSVFMQGEWELLGLVLPFRFSESSSYVYEMKFYV 236
QY 243 TAQRSSMYTAVIYTPASCIYVIALSAFMLPPHNGEKIMINGILLIYIAFLMYFAQL 302
Db 237 VIRRRPLFYVVSLLPFIPLVMQIVGFYLPNN-SGERSFKITLLGLYSVFLLIVSDTL 295
QY 303 PVLNNPPLVYVIFSTSLYISVSTIYEVLVLYATKHKRRRL-----PEALKRLHHTLG 358
Db 296 PATVIGTPLLGVFVVCMLALIVLSLAETIFVRLV---HKQDLQQPVPAWMLRLVLETRI- 351
QY 359 TWLLSVFTTGE-----SOAEKTRK-----MDEH-----PYEADEQESSP----- 395
Db 352 AMLLCLEQGSISQRPATSOATKTDDCSAMGNHSGGQDPFSPKSRDRCSPPPPREA 411
QY 396 -LGINHTVEPGKANQF-----DWALLATAVDRIISFVSFLAFLILAI 437
Db 412 SLAVCGLLQELLSIRQFLERKDEIREVARDMLRVGSYLDKLLFHTLLAVLAYSI 466

RESULT 15
ACHO_RAT STANDARD; PRT; 464 AA.
ID ACHO_RAT
AC P12391;
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neuronal acetylcholine receptor protein, beta-3 chain precursor.
GN CHRN3 OR ACRB3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89197926; Pubmed=2703489;
RA Deneris E.S., Boulter J., Swanson L.W., Patrick J., Heinemann S.F.;
RT "beta 3: a new member of nicotinic acetylcholine receptor gene family
is expressed in brain.";
RT J. Biol. Chem. 264:6268-6272(1989).

```



CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN  
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND  
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA  
CC MEMBRANE.  
CC -1- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT  
CC TYPES OF SUBUNITS: ALPHA AND BETA.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; J04636; AAC28887.1; .  
DR PIR; A33523; A33523.  
DR InterPro; IPR000188; GABAA\_receptor.  
DR InterPro; IPR001175; Neur\_channel.  
DR Pfam; PF02931; Neur\_chan\_LBD; 1.  
DR Pfam; PF02932; Neur\_chan\_memb; 1.  
DR PRINTS; PR00252; NRIONCHANNEL.  
DR PROSITE; PS00236; NEUROTR\_ION\_CHANNEL; 1.  
KW Postsynaptic membrane; Ionic channel; glycoprotein; signal;  
KW Transmembrane; Multigene family.  
FT SIGNAL 1 30 POTENTIAL.  
FT CHAIN 31 464 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,  
FT DOMAIN 31 238 BETA-3 CHAIN.  
FT TRANSMEM 239 263 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 271 288 POTENTIAL.  
FT TRANSMEM 305 326 POTENTIAL.  
FT DOMAIN 327 434 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 435 453 POTENTIAL.  
FT DISULFID 159 173 BY SIMILARITY.  
FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SO SEQUENCE 464 AA; 53261 MW; 5D3B1F7DB934D1D8 CRC64;

Query Match 17.5%; Score 403.5; DB 1; Length 464;  
Best Local Similarity 24.9%; Pred. No. 5.2e-25;  
Matches 121; Conservative 81; Mismatches 178; Indels 105; Gaps 17;

QY 18 LLLQMLMGLMGLTSGVATATADPKNANKALDRHAGLFTNYDSVQPYFGPTNV 77  
DB 9 LVLSTLGSWVTLTATAGLSSVAEHE-----DALRHLFOGYQKWVRPLN----- 55  
QY 78 SLEWVVT-----IDDELNGKLTTHCWLNLWRDEERYWQPSOYDNTQITLKSS 129  
DB 56 SSDIIVFGLKISQLVVDYDNEKNQMTTVMWKQEMTDOKLRWNPDEYGGINSIKVSES 115  
QY 130 VWTPOITLF-NGD--EGGLAETQVTLSDHGFRRMPRAVYATACELMLNMPHDKQSC 185  
DB 116 LMLPDIIVLFENADGRFESSLM--TKAIYKSSGTVSMTPRASYKSSCTMDVTFEPRQNC 173  
QY 186 KLIKISW--GLKVVLPENGARGESLDHDLVQSPENEIYDS-----RAHEVSQDY 235  
DB 174 SMKFGSMYDGTWDL-----LLINENYDRKDFDNGEWEILLAKGKGNRRREGFYSS--Y 226  
QY 236 GMEYTLTAORSSMYTAVITPACIVILALSAFRLPRHMGGEKIMINGILLIYIAFL 295  
DB 227 PFVITYSEVLRLPLFTFLIIPCLGLSFLTVLVYLPD--EGEKLSTSVLVSLTVFL 285  
QY 296 MYFAQLLEVLNNPPLVVIYFYSTLSLVSSTIVEVLVLYLATGKHRRLEPA--LRKL 353  
DB 286 LVIEIIPSSKVIPLIGELYLFIMIFVTLISIVVFINVHNRSSSTYHPRAPVYKRLF 345  
QY 354 HGHLCITWLLISVFTTGESQAETKEMDEHPREADEQESSPLGINHTEVPG----- 405  
DB 346 LQRLPRMLCKM-----DPMDFSPDCKESDITAVRG---KVSGRKQTPAS 388

QY 406 -----AKANQF-----DMALATAVDRIISFVSFSLAFILIA 436  
DB 389 DGERVLVAFLEKASISIVYISRHVYKKEHFISQVYQDMKFFVAQVDRIT-----FLMLFLIAS 444  
QY 437 IRCSV 441  
DB 445 VLGSI 449

Search completed: July 18, 2002, 10:26:53  
Job time: 601 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: July 18, 2002, 09:14:42 ; Search time 69.47 Seconds  
(without alignments)  
609.981 Million cell updates/sec

Title: US-09-732-680A-2

Sequence: 1 MTTTPKIKAPVSGPLPLLL.....RISFVSFLAFLILAIRCSV 441

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR\_71:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	444.5	19.2	528	1	ACCH2N
2	427	18.5	457	1	ACBOAI
3	424	18.3	511	2	A40110
4	418	18.1	627	2	JC4021
5	414.5	17.9	457	2	S13872
6	412.5	17.8	457	2	A24383
7	410	17.7	455	2	A55972
8	410	17.7	455	2	S51116
9	409.5	17.7	625	2	A26456
10	407.5	17.6	445	2	A19458
11	407.5	17.6	622	1	ACCH4N
12	407	17.6	557	2	S12359
13	406	17.6	494	2	T09289
14	405.5	17.5	457	1	ACHUAI
15	403.5	17.5	464	2	A33523
16	400.5	17.3	457	2	A28529
17	398.5	17.2	457	2	S08162
18	397	17.2	479	2	A55382
19	397	17.2	505	2	S07227
20	394.5	17.1	468	2	A38223
21	394.5	17.1	500	2	S12899
22	391.5	16.9	512	2	B37014
23	391	16.9	454	2	B39218
24	391	16.9	454	2	T24695
25	390.5	16.9	456	1	ACCH4N
26	390	16.9	493	2	S68587
27	389.5	16.9	461	2	S15048
28	389	16.8	461	2	T25671
29	389	16.8	523	2	T22910

30	384.5	16.6	501	2	A25338	nicotinic acetylch
31	383.5	16.6	499	2	A24572	nicotinic acetylch
32	378.5	16.4	461	1	ACRYAI	nicotinic acetylch
33	377.5	16.3	501	2	S04607	nicotinic acetylch
34	377	16.3	466	2	S16333	nicotinic acetylch
35	377	16.3	502	1	ACHUAI	nicotinic acetylch
36	377	16.3	502	2	A57175	nicotinic acetylch
37	376	16.3	462	2	S06893	nicotinic acetylch
38	374.5	16.2	542	2	T19862	hypothetical prote
39	372	16.1	501	2	S13873	nicotinic acetylch
40	372	16.1	502	2	G02259	alpha 7 neuronal n
41	371.5	16.1	493	1	ACRYAI	nicotinic acetylch
42	371	16.1	514	1	ACCHG1	nicotinic acetylch
43	369	16.0	423	2	T18056	nicotinic acetylch
44	369	16.0	502	2	T01378	nicotinic receptor
45	368	15.9	503	2	A53956	nicotinic acetylch

#### ALIGNMENTS

RESULT 1  
ACCH2N  
nicotinic acetylcholine receptor alpha-2 chain precursor, neuronal - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 21-Jul-2000  
C:Accession: S00377  
R:Net, P.: Oneyear, C.: Alliod, C.: Couturier, S.: Ballivet, M.  
EMBO J. 7, 595-601, 1988  
A:Title: Genes expressed in the brain define three distinct neuronal nicotinic acetylcholine receptor subunit families  
A:Reference number: S00376; MUID:88283624  
A:Accession: S00377  
A:Molecule type: DNA  
A:Residues: 1-528 <NER>  
A:Cross-references: EMBL:X07339; NID:962792; PIDN:CAB59645.1; PID:96136914  
C:Genetics:  
A:Introns: 21/1; 73/3; 88/3; 125/2; 487/3  
C:Superfamily: acetylcholine receptor  
C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membra  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-528/Product: nicotinic acetylcholine receptor alpha-2 chain #status predicted <M1>  
F:241-264/Domain: transmembrane #status predicted <TM1>  
F:272-290/Domain: transmembrane #status predicted <TM2>  
F:306-327/Domain: transmembrane #status predicted <TM3>  
F:502-520/Domain: transmembrane #status predicted <TM4>  
F:54,104/Binding site: carbohydrate (asn) (covalent) #status predicted  
F:158-172,222-223/Disulfide bonds: #status predicted

Query Match 19.2% Score 444.5; DB 1; Length 528;  
Best Local Similarity 27.7% Pred. No. 1.7e-28;  
Matches 119; Conservative 82; Mismatches 14; Indels 81; Gaps 15;

QY	39	ATADPKANVAKALDRHLAGLFTNYDSVQVPVQGTPTNLSLMVYVY-----IDIDEL	91
DB	22	ATREQKQHPGFEDRLFKHLFTGYNKMSRP-----PNTSDVVIYKFGSLNQILDVDEK	76
QY	92	NKRLTHCWLNRMRDEEVWQPSQYDNTQTTLASSEVWTOITLFFNDEGLAEIOV	151
DB	77	NOMMTTNVWLKQWSDYKLRNMPDEDFDNTSIRVSEMIWIPDIYVNNADGE-FAVTAM	135
QY	152	TLSSH---DGFRMPMPAVYATVACELNMLNMPDKOSCKLKISW--GLKVLPENGTAR	205
DB	136	TRAHLEFSNGKVKWVPDAIKSSCSIDVYFFPDQCKMKFESWYDKAKIDL-----	188
QY	206	GSGLDH---DDLVOSEMEIVDSRAHFVSQDY-----YGYMEYTLTAQRSSMTAVI	255
DB	189	ENMEHHVDLKYWMSGEGEAIINAGRYNSKKYDCOTEIPIPTFFVIRLPLFTIIL	247
QY	256	YRPASCTVTLASAFMLPPLMGCEKIMGLIYIAAFKFAQLPLVLSNNPLVYIF	315
DB	248	IPCLILISCTVIVLPFLPDC-GEKITLICISVLSLTVFLTLITETIPSTSLVPIIGEX	306

```

QY 316 YSTSLYSLVSTVEVLVLYLATGKHKR-----RLPEALRKLHGLGTV-----LLL 363
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 307 LLEFTMTFTSLITITVEVLNV-----HHRSPSTHMPHVSFFGLFIPRMFLMKRPPLL 362
QY 364 SVFSTTG-----ESQAEKTKEMDEHDEHYEADQESSPLGINTEV 403
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 363 PAECTGTQYPPPGTRTSLRSCWLETVDVDRKWEHEEEEEEEEEEEERKAPV-----SRV 417
QY 404 P--GAKANQ 410
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 418 PGGSGSGCTQ 426

```

## RESULT 2

```

ACB0A1
C:Species: Bos primigenius taurus (cattle)
C>Date: 18-Apr-1984 #sequence_revision 18-Apr-1984 #text_change 22-Jun-1999
C:Accession: A03169
R:Noda, M.; Furutani, Y.; Takahashi, H.; Toyosato, M.; Tanabe, T.; Shimizu, S.; Kikyoan
Nature 305, 818-823, 1983
A:Title: Cloning and sequence analysis of calf cDNA and human genomic DNA encoding alpha
A:Reference number: A03168; MUID:84039794
A:Accession: A03169
A:Molecule type: mRNA
A:Residues: 1-457 <NOD>
A:Cross-references: GB:X02509; NID:949; PIDN:CAA26345.1; PID:950
A>Note: four hydrophobic transmembrane segments are found in each of the four kinds of c
-318, and 429-447 in the alpha chain
C:Comment: The functional receptor molecule has two alpha chains and one each of the bet
C:Superfamily: acetylcholine receptor
C:Keywords: glycoprotein; ion channel; membrane protein; neurotransmitter receptor; post
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-457/Product: nicotinic acetylcholine receptor alpha chain #status predicted <MAT>
F:148-162/Disulfide bonds: #status predicted
F:161/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

Query Match 18.5%: Score 427; DB 1; Length 457;

Best Local Similarity 27.4%: Pred. No. 3,9e-27;

Matches 129; Conservative 91; Mismatches 163; Indels 88; Gaps 22;

```

QY 17 PLTIOMIMGLMGLTSPGATATADPRKANVKAALDRHAGLFNTYDSVQPV---FGGT 73
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 PLTL--LLGLCSAGL--VLGSEHET-----RLVAKLFEDYNSVAVRVDHQAQAV 49
QY 74 PTNVSLEMYVTYIDDELNGKLTTHCWLNRWDEERYWQSOYDNTQITLKSSEVTP 133
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 50 EYVYGLDQ--IQILNDEVNOIYITNVRKQOWDYNLKNRPDYGCVAKIHIPEKTIWR 108
QY 134 QITLFNGDEG--GIMAEFOVTLSDHGFRMRPRAVYTAECELNMLNPHDKOSCKRLIGS 191
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 109 DLVLYNNAADGFAIVKFTKVLVDYTGHTWTPRAIFKSYCEIITVHPRPDQNSMKLGT 168
QY 192 W-----GLKVVLPENGTRAGESLDHDDLVSPEEIVDSRA--HYVSODYIG-----YME 239
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 169 WYDGSYVYVNPES-----DQPDLSNEMSEGEVYIKESRGKHWV---FYACSPSTYLD 220
QY 240 --TTLAQRSSMYTAVIYTPASCTIVILASFWLPRHMGSEKIMINGLLIYIAAFIMY 297
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 221 ITHFWNQRLPYFTIVVITPCLLFSLTGLVFTLPD--SEKMTLSIVLSLFTVFLV 279
QY 298 PAQLPLVLSNNTPLVLIYFYSTLSLYSTVEVLVLYLATGKHKR-----LPEALRKL 352
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 280 IVELIPTSSAVPLIGKVMFTWVFYIASIITVIT-----NTHHRSPTTHMPEWVRKY 335
QY 353 LHGLGLTWLLISVSTGESQAEKTKEMDEHY--EADQESS-----PLG-----INH 400
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 336 F---IDTIPIMFSTMKRPSREK---QDKKIFTEIDIDISIDSGKPPRPMGFGHSPILKH 389
QY 401 TEVPGA-----KANO-----FDWALLTAVDRIISFVSFLAFLI 434
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 390 PEKASALEGIKYIAETMKSDQESNNAAEWVKYVAMVDHILLAVFMLVCI 440

```

## RESULT 3

nicotinic acetylcholine receptor alpha-2 chain precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 20-Mar-1992 #sequence\_revision 10-Apr-1992 #text\_change 20-Aug-1999

C:Accession: A40110

R:Wada, K.; Ballivet, M.; Boulter, J.; Connolly, J.; Wada, E.; Denieris, E.S.; Swanson

Science 240, 330-334, 1988

A:Title: Functional expression of a new pharmacological subtype of brain nicotinic ac

A:Reference number: A40110; MUID:88178113

A:Accession: A40110

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-511 <NMD>

A:Cross-references: GB:M20297; NID:g202672; PIDN:AAA40664.1; PID:g202674

A>Note: the authors translated the codon TCG for residue 494 as Cys

C:Superfamily: acetylcholine receptor

C:Keywords: neurotransmitter receptor

Query Match 18.3%: Score 424; DB 2; Length 511;

Best Local Similarity 24.4%: Pred. No. 8e-27;

Matches 126; Conservative 95; Mismatches 167; Indels 128; Gaps 18;

```

QY 20 LQIMMGLMGLTSPGATATADPRKANVKAALDRHAGLFNTYDSVQPVFGGTPTNVS 79
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 8 LQFTNHLXWMLLVF---AVLTQGSHTNEDRLFKHLFGYIRMRAPV-----PMTSD 59
QY 80 EMVVTY-----IDDELNGKLTTHCWLNRWDEERYWQSOYDNTQITLKSSEVWT 132
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 VVIVRFGISIAQLDVEDENKQMTNTNWLKQEMNDYKLRMPAEFGVTSIRVSEMIWI 119
QY 133 PQLTFNGDEGMAEQVTLTSH---DGHFRMRPRAVYTAECELNMLNPHDKOSCKRLI 189
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 PDIVLYNNADEG--FVTHMTKAHLFTGYVHWVPRATKSCSIDVFFFPDQNCMKRF 178
QY 190 GSW---GLKVVLPENGTRAGESLDHDDLVSPEEIVDSRAHFVSODY-----YGYMEY 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 GSWIYDRAKALIDEG---MERTVDLKDYMSEGAIIINAGTYSKKYDCAELIYRPVY 234
QY 241 TLTAQRSSMYTAVIYTPASCTIVILASFWLPRHMGSEKIMINGLLIYIAAFIMY 300
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 235 YFVIRRLRPLEFTYINLIIPCLLSCTIVLVFYLPEEC--GEKTLTCLISVLSLFTVFLITE 293
QY 301 LLPLVLSNNTPLVLIYFYSTLSLYSTVEVLVLYLATGKHKR-----RLPEALRKLHG 355
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 294 IIPSTSLVPLIGELYLFTMTLITSLIYIVFVLNV-----HHRSPSTHMPHVRVALLG 349
QY 356 HLGTWLLS-----VFSTGESQAEKTKEMDEHPEADQESS 393
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 350 RVRPRLMNNRPLRPMELHGPSDLKLSYHMLFTNMDAGEEETEE-----EEEEEDN 403
QY 394 -----SPLGINH-----TEVPAKANQF----- 411
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 404 ICVCAGLPDSSMGVLYGHGHLRAMPBEKTP--SQASEILSPQIOKALEGVHIADRL 462
QY 412 -----DWALLTAVDRIISFVSFLAFLI 436
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 463 RSEDADSSVKEDMKYVAMVDRI---FLWLFITVS 494

```

## RESULT 4

nicotinic acetylcholine receptor alpha-4 chain, neuronal - human

C:Species: Homo sapiens (man)

C>Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 20-Aug-1999

C:Accession: J04021; S55471

R:Monteggia, L.M.; Gopalakrishnan, M.; Touma, E.; Idler, K.B.; Nash, N.; Arneric, S.P

Gene 155, 189-193, 1995

A:Title: Cloning and transient expression of genes encoding the human alpha 4 and bet

A:Reference number: J04021; MUID:95237608



```

Db      289 SAVPLIGKYLFTWVFAVSLITIVYI---NTHNRSPSTHMEVWRKVFIDIPNIM 344
      362 LLSVSTTGESQAEK---TKMEDHPYEADQESSPLG-----INHEVGA----- 406
      345 FFSFMKRPSSNRKQKRIETEDIDS--DISGKPPRPGRFSPPLIKHEVSAIIEGVYI 402
Qy      407 ----KANO-----FDMLATAVDRISFVSFLAFLI 434
      403 AETMKSDQESNNAAEWKYVAMVDHILLGVFMLVCLL 440

```

## RESULT 7

nicotinic acetylcholine receptor beta-3 chain precursor - chicken

C:Species: Gallus gallus (chicken)  
C:Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 21-Jul-2000

C:Accession: A55972  
R:Hernandez, M.C.; Erkman, L.; Matter-Sadzinski, L.; Roztocil, T.; Ballivet, M.; Matter, J. Biol. Chem. 270, 3224-3233, 1995

A:Title: Characterization of the nicotinic acetylcholine receptor beta3 gene. Its regula  
A:Reference number: A55972; MUID:95155414

A:Accession: A55972  
A>Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-455 <HER>

A:Cross-references: GB:X83739; NID:96165256; PIDN:CA59814.1; PID:96165257

C:Superfamily: acetylcholine receptor  
C:Keywords: neurotransmitter receptor

Query Match 17.7%; Score 410; DB 2; Length 455;  
Best Local Similarity 26.3%; Pred. No. 9.7e-26;  
Matches 115; Conservative 85; Mismatches 167; Indels 70; Gaps 17;

```

Qy      52 DRLHAGFTNYDSVDPVFOGTPT-NVSLKMYVT-YIDIDELNCKLTTHCWLNRWDEE 109
      27 DALLRHLEFGYQKWRVPEVNSNDTIKVLFGKISQLDVDEKQMLMTTNWLRKQEMMDHK 86
Qy      110 RVWQPSQYDNTIQTTLKSSEVWTPQITLF-NGD---EGGLMAETQVTLSHDGRFRMPRA 165
      87 LSNWPEEYGGITAIRVSESLMLPDIVLFENADGRFEGSLM--TKAIVKYNQVQWMPRA 144
Qy      166 VYTAAYCELNLNMPHDKOSCKLKGSGW----LKVVLPENGTARGESLDHDLVQSPDW 220
      145 SYKSSCTMELTFPFDDQNCMKFGSWTYDGSWVDLIVD-----ENVDTKDFPDNGEW 198
Qy      221 EYDSRAHFVSQD---YGYMEYTLTAQRSSMYTAVITYTPASCIYIILASAFMLPPHM 276
      199 EILNAKMKGNRKDGILSYFPVTYSFVLRRLPLFYTLFLIPLCLGSLFLVLFVLEPLSD- 257
Qy      277 GGEKIMINGLLIYIAFLMYFAQLLPVLSNMPLVVIFYSTSLYSVSTIVEVLVLYL 336
      258 EGEKLSLSTVLSVLFVLELIEIIPSSSKVIRPLIGEYLFIMIFVTLSTIVVFIYINV 317
Qy      337 ATGKHNR-----RLEPALRKLLHGLGTWLLSVFSTTGESQAE-KT 377
      318 -----HHRSSATYHPMAVWRKLFQKLPRL--CMKGHVDRY-----SFSDEEEKETTLKS 367
Qy      378 KEMDEHPYEADQESSPLGINHTEVPG-----AKANOF-----DMALLATAVDRIS 424
      368 KLPGKQKHQAKQDEKVIYAFLEKADSIYIRSHVKKDAFIROYVDWKFVAAQVDRI- 426
Qy      425 FVSFSLAFLILAIRCSV 441
      427 ---FLMLFLVAVTGSV 440

```

## RESULT 8

nicotinic acetylcholine receptor (nAChR) beta-3 chain precursor - chicken

C:Species: Gallus gallus (chicken)

C:Date: 07-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 20-Aug-1999

C:Accession: S51116  
R:Hernandez, M.C.; Erkman, L.; Matter-Sadzinski, L.; Roztocil, T.; Ballivet, M.; Matter  
Submitted to the EMBL Data Library, January 1995  
A:Description: Characterization of the nicotinic acetylcholine receptor beta-3 gene.  
A:Reference number: S51116  
A:Accession: S51116  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-455 <HER>

A:Cross-references: EMBL:X83739; NID:9633182; PIDN:CA58707.1; PID:9633183

C:Superfamily: acetylcholine receptor  
C:Keywords: neurotransmitter receptor

Query Match 17.7%; Score 410; DB 2; Length 455;  
Best Local Similarity 26.3%; Pred. No. 9.7e-26;  
Matches 115; Conservative 85; Mismatches 167; Indels 70; Gaps 17;

```

Qy      52 DRLHAGFTNYDSVDPVFOGTPT-NVSLKMYVT-YIDIDELNCKLTTHCWLNRWDEE 109
      27 DALLRHLEFGYQKWRVPEVNSNDTIKVLFGKISQLDVDEKQMLMTTNWLRKQEMMDHK 86
Qy      110 RVWQPSQYDNTIQTTLKSSEVWTPQITLF-NGD---EGGLMAETQVTLSHDGRFRMPRA 165
      87 LSNWPEEYGGITAIRVSESLMLPDIVLFENADGRFEGSLM--TKAIVKYNQVQWMPRA 144
Qy      166 VYTAAYCELNLNMPHDKOSCKLKGSGW----LKVVLPENGTARGESLDHDLVQSPDW 220
      145 SYKSSCTMELTFPFDDQNCMKFGSWTYDGSWVDLIVD-----ENVDTKDFPDNGEW 198
Qy      221 EYDSRAHFVSQD---YGYMEYTLTAQRSSMYTAVITYTPASCIYIILASAFMLPPHM 276
      199 EILNAKMKGNRKDGILSYFPVTYSFVLRRLPLFYTLFLIPLCLGSLFLVLFVLEPLSD- 257
Qy      277 GGEKIMINGLLIYIAFLMYFAQLLPVLSNMPLVVIFYSTSLYSVSTIVEVLVLYL 336
      258 EGEKLSLSTVLSVLFVLELIEIIPSSSKVIRPLIGEYLFIMIFVTLSTIVVFIYINV 317
Qy      337 ATGKHNR-----RLEPALRKLLHGLGTWLLSVFSTTGESQAE-KT 377
      318 -----HHRSSATYHPMAVWRKLFQKLPRL--CMKGHVDRY-----SFSDEEEKETTLKS 367
Qy      378 KEMDEHPYEADQESSPLGINHTEVPG-----AKANOF-----DMALLATAVDRIS 424
      368 KLPGKQKHQAKQDEKVIYAFLEKADSIYIRSHVKKDAFIROYVDWKFVAAQVDRI- 426
Qy      425 FVSFSLAFLILAIRCSV 441
      427 ---FLMLFLVAVTGSV 440

```

## RESULT 9

A26456

nicotinic acetylcholine receptor alpha-4 chain, brain precursor - rat (fragment)  
N:Alternate names: nicotinic acetylcholine receptor acetylcholine-binding chain; nico  
C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 13-Nov-1998

C:Accession: A26456; S02415

R:Goldman, D.; Deneris, E.; Luyten, W.; Kocher, A.; Patrick, J.; Heinemann, S.  
Cell 48, 965-973, 1987

A:Title: Members of a nicotinic acetylcholine receptor gene family are expressed in d  
A:Reference number: A26456; MUID:87159533

A:Accession: A26456  
A:Molecule type: mRNA  
A:Residues: 1-625 <GOLD>

A:Cross-references: EMBL:M15681; NID:9205637; PID:9205638  
R:Whiting, P.; Esch, F.; Shimazaki, S.; Lindstrom, J.  
FEBS Lett. 219, 459-463, 1987

A:Title: Neuronal nicotinic acetylcholine receptor beta-subunit is coded for by the c  
A:Reference number: S02415; MUID:87276531

A:Accession: S02415  
A:Molecule type: protein  
A:Residues: 'X', 28-30, 'X', 32, 'X', 34-36, 'X', 38-40, 'X', 42-43 <WHIT>

us-09-732-680a-2.rpt

Dy 168 TAYCELNMLNPDKOSCKLIGSW---GLKVVLPENGSTARGESLDHDDLVGSEMEIV 223  
::: : :: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 133 KSYCIITVTHPEQONCSMLGIWTYDGSVAALNPES---DQPLSNMGEWEVIK 187

Db 188 EARGKHWV--FYSCPTPIYIDITYHFWMORLPFLFYIVNVIIPCLLFSSELTSLVHYLP 244

```

      : | :| :| : | :| :|
Db 304 I-----NTHHSPSTHIMPEWVKVFIIDITPIMFESTMKRPSRDQOEKRIFEDIDIS-- 357

```

```

: : | | | | : : : : : :
Db 358 DISCRPPPGFHSPLIKHPEVKSALCEVKYIAETMKSQESNNAAEEFKYVAMVDHI 417

```

Db 418 LIGVEMLVCLI 428

RESULT 11  
ACCH4N  
nicotinic acetylcholine receptor alpha-4 chain precursor, neuronal - chick

CiDate: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 21-Jul-2000  
CiAccession: S00379; A38756; A61186  
RiNef, P.; Oneyser, C.; Alliod, C.; Couturier, S.; Ballivet, M.

A:Title: Genes expressed in the brain define three distinct neuronal niches  
A:Reference number: S00376; MUID:88283624  
A:Accession: S00379

A:Residues: 1-622 <NEF>  
A:Cross-references: GB:X07348; GB:Y00834; NID:g62809; PIDN:CAA30285.1; PIII  
A:Accession: A38756

A:Residues: 21-622 <NE2>  
A:Cross-references: EMBL:X07348  
R:Whiting, P.J.; Schoepfer, R.; Conroy, W.G.; Gore, M.J.; Keyser, K.T.; Sh

A:Title: Expression of nicotinic acetylcholine receptor subtypes in brain  
A:Reference number: A61186; MUID:9127868  
A:Accession: A61186

A:Residues: 24-29, 'XX', 32 <WHT>  
C:Genetics:  
A:Introns: 21/1; 71/3; 86/3; 123/2; 581/3  
C:Surrounding: Acetylcholine receptor

C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-622/Product: nicotinic acetylcholine receptor alpha-4 chain #status e  
F:230-260/Domain: transmembrane #status predicted <tmh>

Accession	Domain	status	predicted	TM2
F.270-288	Domain	#status	predicted	<TM2>
F.304-325	Domain	#status	predicted	<TM3>
F.596-614	Domain	#status	predicted	<TM4>

F:156-170, 220-221/Disulfide bonds: #status predicted

50. AIDPIHAGT EENVDSNDVODFOGCTBPNVSI EMVUYEV-----IDPDEI NCTI EEUHGGI N 103  
Best Local Similarity -29.4%; Pred. NO. 2.4e-25;  
Matches 96; Conservative 73; Mismatches 119; Indels 39; Gaps

```

Db      31 AEERLKKLFGYNNKSRPV-----ANISDVVLVRFGLSLAQLLIDVDEKNQMMTTNWVK 85

```

Dy 168 TAYCELNMLNPDKOSCKLIGSW---GLKVVLPENGSTARGESLDHDDLVGSEMEIV 223  
::: : :: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 133 KSYCIITVTHPEQONCSMNLGTWYDGSVVAINPES---DQPLSNMEGEVNIK 187

Db 188 EARGKHWV--FYSCPTPIYIDITYHFWMORLPFLFIYNVIIPCLLFSSELTSLVEYLP 244

```

      ||: :: :: : ||: :::| | : ||: : ::: | | : ||
Db 245 TD-SGKMTLSISVLLSTVTFELVIVELIPSTSSAVPLGKYMFLTWVFVIAIIITIV 303

```

```

      : | :| :| : | :| :|
Db 304 I-----NTHHSPSTHIMPEWVKVFIIDITPIMFESTMKRPSRDQOEKRIFEDIDIS-- 357

```

```

      : : | : | | | : : : : : : : :
Db 358 DISCRPPPGFHSPLIKHPEVKSALIECVKYIAETMKSQESNNAAEEFKYVAMVDHI 417

```

Db 418 LTGVMLVCLI 428

RESULT 11  
ACCH4N  
nicotinic acetylcholine receptor alpha-4 chain precursor, neuronal - chick

CiDate: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 21-Jul-2000  
CiAccession: S00379; A38756; A61186  
RiNef, P.; Oneyser, C.; Alliod, C.; Couturier, S.; Ballivet, M.

A:Title: Genes expressed in the brain define three distinct neuronal nicotinic receptor subtypes  
A:Reference number: S00376; MUID:88283624  
A:Accession: S00379

A:Residues: 1-622 <NEF>  
A:Cross-references: GB:X07348; GB:Y00834; NID:g62809; PIDN:CAA30285.1; PIII  
A:Accession: A38756

A:Residues: 21-622 <NE2>  
A:Cross-references: EMBL:X07348  
R:Whiting, P.J.; Schoepfer, R.; Conroy, W.G.; Gore, M.J.; Keyser, K.T.; Sh

A:Title: Expression of nicotinic acetylcholine receptor subtypes in brain  
A:Reference number: A61186; MUID:9127868  
A:Accession: A61186

A:Residues: 24-29, 'XX', 32 <WHT>  
C:Genetics:  
A:Introns: 21/1; 71/3; 86/3; 123/2; 581/3  
C:Surrounding: Acetylcholine receptor

C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-622/Product: nicotinic acetylcholine receptor alpha-4 chain #status e  
F:230-260/Domain: transmembrane #status predicted <tmh>

Accession	Protein	Domain	status	predicted	TM2
F.270-288	transmembrane	#status predicted	<TM2>		
F.304-325	transmembrane	#status predicted	<TM3>		
F.596-614	transmembrane	#status predicted	<TM4>		

F:156-170, 220-221/Disulfide bonds: #status predicted

Best Local Similarity 29.4%; Pred. NO. 2.4e-25;  
Matches 96; Conservative 73; Mismatches 119; Indels 39; Gaps

```

Db      31 AEERLKKLFGYNGKSRPV-----ANISDVLVRFGLSLAQQLIDVDEKNGQMTTNVWK 85

```

[illegible][illegible]





Fri Jul 19 08:26:22 2002

us-09-732-680a-2.rpr

Page 8

Search completed: July 18, 2002, 10:19:32  
Job time: 3890 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 08:54:47 ; Search time 66.67 seconds  
(without alignments)  
161.567 Million cell updates/sec

Title: US-09-732-680A-2  
Perfect score: 2311  
Sequence: 1 MTTTPKIKAPVSGPLPLTL.....RISFVSFLAILAIRCSV 441

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PTOUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/Backfilest.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	423.5	18.3	629	US-08-464-258B-6	Sequence 6, Appli
2	423	18.3	511	US-08-464-258B-4	Sequence 4, Appli
3	422.5	18.3	629	US-08-278-635B-6	Sequence 6, Appli
4	422.5	18.3	629	US-08-471-961-6	Sequence 6, Appli
5	418.5	18.1	510	US-08-278-635B-4	Sequence 4, Appli
6	418.5	18.1	510	US-08-471-961-4	Sequence 4, Appli
7	416	18.0	529	US-08-496-855A-2	Sequence 2, Appli
8	414.5	17.9	457	US-08-278-635B-3	Sequence 3, Appli
9	414.5	17.9	457	US-08-471-961-3	Sequence 3, Appli
10	414	17.9	458	US-08-464-258B-3	Sequence 3, Appli
11	412.5	17.8	528	US-08-466-589-2	Sequence 2, Appli
12	412.5	17.8	528	US-08-700-636-2	Sequence 2, Appli
13	412.5	17.8	528	US-08-467-574-2	Sequence 2, Appli
14	412.5	17.8	528	US-09-217-345-2	Sequence 2, Appli
15	397	17.2	479	US-08-278-635B-2	Sequence 2, Appli
16	397	17.2	479	US-08-464-258B-2	Sequence 2, Appli
17	397	17.2	479	US-08-471-961-2	Sequence 2, Appli
18	384	16.6	627	US-08-466-589-6	Sequence 6, Appli
19	384	16.6	627	US-08-700-636-6	Sequence 6, Appli
20	384	16.6	627	US-08-467-574-6	Sequence 6, Appli
21	384	16.6	627	US-09-217-345-6	Sequence 6, Appli
22	381.5	16.5	449	5468481-1	Patent No. 5468481
23	381	16.5	502	US-08-278-635B-7	Sequence 7, Appli
24	381	16.5	502	US-08-464-258B-7	Sequence 7, Appli
25	381	16.5	502	US-08-471-961-7	Sequence 7, Appli
26	380.5	16.5	497	US-08-278-635B-5	Sequence 5, Appli
27	380.5	16.5	497	US-08-464-258B-5	Sequence 5, Appli

28	380.5	16.5	497	3	US-08-471-961-5	Sequence 5, Appli
29	373	16.1	502	2	US-08-466-589-8	Sequence 8, Appli
30	373	16.1	502	2	US-08-700-636-8	Sequence 8, Appli
31	373	16.1	502	3	US-08-467-574-8	Sequence 8, Appli
32	373	16.1	502	4	US-09-217-345-8	Sequence 8, Appli
33	371.5	16.1	449	3	US-08-462-351-2	Sequence 2, Appli
34	370.5	16.0	504	2	US-08-466-589-4	Sequence 4, Appli
35	370.5	16.0	504	2	US-08-700-636-4	Sequence 4, Appli
36	370.5	16.0	504	3	US-08-467-574-4	Sequence 4, Appli
37	370.5	16.0	504	4	US-09-217-345-4	Sequence 4, Appli
38	360.5	15.6	449	6	5194425-1	Patent No. 5194425
39	356	15.4	502	2	US-08-466-589-10	Sequence 10, Appli
40	356	15.4	502	2	US-08-700-636-10	Sequence 10, Appli
41	356	15.4	502	3	US-08-467-574-10	Sequence 10, Appli
42	356	15.4	502	4	US-09-217-345-10	Sequence 10, Appli
43	354.5	15.3	511	1	US-08-278-635B-8	Sequence 8, Appli
44	354.5	15.3	511	3	US-08-464-258B-8	Sequence 8, Appli
45	354.5	15.3	511	3	US-08-471-961-8	Sequence 8, Appli

## ALIGNMENTS

RESULT 1  
US-08-464-258B-6  
; Sequence 6, Application US/08464258B  
; Patent No. 6013766  
; GENERAL INFORMATION:  
; APPLICANT: ELGOYHEN, ANA BELEN  
; APPLICANT: JOHNSON, DAVID S.  
; APPLICANT: BOULTER, JAMES R.  
; APPLICANT: HEINEMANN, STEPHEN F.  
; TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GRAY CARY WARE & FREIDENRICH LLP  
; STREET: 4365 EXECUTIVE DRIVE, SUITE 1600  
; CITY: SAN DIEGO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 92121  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/464,258B  
; FILING DATE: 06/05/95  
; CLASSIFICATION: 530  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 08/278,635  
; FILING DATE: 21-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: REITER, STEPHEN E.  
; REGISTRATION NUMBER: 31,192  
; TELEPHONE: 619-677-1409  
; TELEFAX: 619-677-1465  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 629 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-464-258B-6

Query Match 18.3%; Score 423.5; DB 3; Length 629;  
Best Local Similarity 26.1%; Pred. No. 5e-36;  
Matches 121; Conservative 83; Mismatches 176; Indels 83; Gaps 13;

```

01 12 SGPGLPRLLOKMLGLTGLTSGCATATAPKNAVKALDRLHAGFTTSDSVOCPVQ 71
02 4 SGPAPRPLLLPLLLLLGLTGLLRASSIETERRAAERLLKRL----LFSGKNKSRV-- 57
03 72 GTPNVLSLEAVVY-----IDIDELNGKLTTGHCWILMWREDERVWQSPQSDNTOIT 124
04 58 ---GNISDVLVLRGLSLIAQLIDVDEKQNMATTVMWVQOEWHNDYKLRMPDPCDYENTVSIR 1144
05 125 LKSSVWVTPOTLLENGDGGGLMATAQVYLSH---DGHFRMMRPVAVYTAUCELMLNMPHD 168
06 115 IPELSIMRPDVLVNNMAGD-FATYHLTKAHLFYDGRVQWMPRAIVSSSIDVTEPFED 1737
07 182 KQSKLKLKGSW---GLKVLVPENGTARGEISDHDLDVQSPEMEIVDSRAHFVSODY---- 234
08 174 QONCMKMGSMVTYDKAKIDL-----VSIHSRDQDLDFMESGSMVYLDVAGVITNPKRYECCA 225
09 235 --YGYMEYTLTAQRSSMYTAIVITPASCIVIALSAFWLPRHNGEKKIMINGILLIVIA 2923
10 230 EIPDYTAFTIIRRLPLFYTTINLLIPCLLISCLTIVLFEYLPSEC-GEKVYCCASSVLLST 2688
11 293 AFIMVFAQLLVLSNPNPLVYVIFSTLSLYLSTVYEVLYLYATGKHK----- 3424
12 289 VFLLTETIIPSTSLVPLDIGEVLFTFIETVLTSLVITVEVLNV---HHSPPRTHTPAW 345
13 343 -----RRLPEALRKLLHG-----HLGTWLLSVFSTTGE 3717
14 346 VRVAVELVPLRLLEPMKRPSVYKDKNCRRILIESMHKMANAPRPREPVEBERGLSDICNOGL 405
15 372 SOAEKTKEMDEHPYEADEQSSPLGINHTVEVPCAKANOEDMA 414
16 406 SPAPTFCNPTDTAVETOPTCRSPPL-----EVPDLKTSSEVEKA 443
17
18 RESULT 2
19 US-08-464-258B-4
20 Sequence 4, Application US/08464258B
21 Patent No. 6013766
22 GENERAL INFORMATION:
23 APPLICANT: ELGOYHEN, ANA BELEN
24 APPLICANT: JOHNSON, DAVID S.
25 APPLICANT: BOULTER, JAMES R.
26 APPLICANT: HEINEMANN, STEPHEN F.
27 TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL *
28 TITLE OF INVENTION: ACETYLCOLINE-GATED ION CHANNEL RECEPTOR
29 NUMBER OF SEQUENCES: 8
30 CORRESPONDENCE ADDRESS:
31 ADDRESSEE: GRAY CARY WARE & FREIDENRICH LLP
32 STREET: 4365 EXECUTIVE DRIVE, SUITE 1600
33 CITY: SAN DIEGO
34 STATE: CALIFORNIA
35 COUNTRY: USA
36 ZIP: 92121
37
38 COMPUTER READABLE FORM:
39 MEDIUM TYPE: Floppy disk
40 COMPUTER: IBM PC compatible
41 OPERATING SYSTEM: PC-DOS/MS-DOS
42 SOFTWARE: Patent In Release #1.0, Version #1.25
43 CURRENT APPLICATION DATA:
44 APPLICATION NUMBER: US/08/464,258B
45 FILING DATE: 06/05/95
46 CLASSIFICATION: 530
47 PRIOR APPLICATION DATA:
48 APPLICATION NUMBER: 08/278,635
49 FILING DATE: 21-JUL-1994
50 ATTORNEY/AGENT INFORMATION:
51 NAME: REITER, STEPHEN E.
52 REGISTRATION NUMBER: 31,192
53 REFERENCE/DOCKET NUMBER: P41 9989
54 TELECOMMUNICATION INFORMATION:
55 TELEPHONE: 619-677-1403
56 TELEFAX: 619-677-1465
57 INFORMATION FOR SEQ ID NO: 4:

```

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 511 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-464-258B-4

Query Match      18.3%; Score 423; DB 3; Length 511;
Best Local Similarity 24.5%; Pred. No. 4e-36;
Matches 126; Conservative 94; Mismatches 167; Indels 128; Gaps 18;

QY 20 LQMLMGMLMGITSPGATATADPRKANVAKALDRLHAGLFTNYSDSDVQVFGPTPNVSL 79
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D8 8 LQFWHTLWLCMLLVP---AVLQGSSTHTHADRLFKHLFGYNNMARVP-----PNTSD 59
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 80 EMVNYI-----IDIDELNGKLTTHCNLNRKREERWQPSOYDNTQITLKSSEVWT 132
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D8 60 VVIATFEGSLIAQLIDVDRKNQMTTNNVLMKDEMNQYKLRMDAERGNVTSLKVPSEMTWT 119
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 133 PQLTTFNDEGGIAMEQVTLSSH---DGHFRMPRAVYATAYCELMMLNMPHDKQSCDKLT 189
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D8 120 PDVLYNNADE-FAVTHMTKAHLFFGTGVHVPRAIKSSSIDVTFPPFPQCKMKKF 178
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 190 GSW---GLKVVLPENGARGESLDHDLVQSPDEWEIFYDSRAHFVSQDY-----YGYMEY 240
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D8 179 GSWYDRAKKIDLEQ---MERTVDLKDWESEGEWAIINATGYNSKKKDYCCAEIYPDVY 234
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 241 TLTQGRSSMTAVIYTPASCIYIALSAFMLPRIMGGEKTIINGLLIIVIAAFILMPYPAQ 300
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D8 235 YFVIRRLDPLFTINILIPCLLSCLTVLYFLPSEC-GEKITLCSVLSLTVFLILITE 293
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 301 LLPVLNNTPPLVIFYSTLSLYLSVSTIVENVLYLYLATGKHNR-----RLPEALRKLHG 355
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D8 294 IIPSTSLVPLIGELYLFETMIFVLSIYIVFVLNV-----HHRSPSTNNMPWYRVALLG 349
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 356 HLGTVLLLS-----VFSTGGSQAQKTEKEMDBHREEDBQES 393
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D8 350 RVPRLMNNRPLPRNELHGSPLDKLSPSYHMLETMMDAGEREYTE-----EEEBEDEN 403
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 394 -----SPLGINH-----TEVPKAKANQF----- 411
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D8 404 ICVCAGLRDSSMGVLYHGGLHRLAMEDETKTP-SQASEILLSPQIQALBGSVHYIADRL 462
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 412 -----DWALLATAVDRISVSSSLAFLIL 435
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D8 463 RSEDADSSVKEDWKYVAMVVDRI-----FLMLEFIIV 493
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT      3
US-08-278-635B-6
: Sequence 6, Application US/08278635B
: Patent No. 5683912
: GENERAL INFORMATION:
: APPLICANT: ELGOYHEN, ANA BELEN
: APPLICANT: JOHNSON, DAVID S.
: APPLICANT: BOULTER, JAMES R.
: APPLICANT: HEINDMANN, STEPHEN F.
: TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL
: TITLE OF INVENTION: ACETYLCHOLINE-GATED ION CHANNEL RECEPTOR
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GRAY CARY WARE & FREIDENRICH
: STREET: 4365 EXECUTIVE DRIVE, SUITE 1600
: CITY: SAN DIEGO
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 92121
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0 Version #1.25

```



```

? APPLICANT: ELGOYHEN, ANA BELEN
? APPLICANT: JOHNSON, DAVID S.
? APPLICANT: BOULTER, JAMES R.
? APPLICANT: HEINEMANN, STEPHEN F.
? TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL
? TITLE OF INVENTION: ACETYLCHOLINE-GATED ION CHANNEL RECEPTOR
? NUMBER OF SEQUENCES: 8
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: GRAY CARY WARE & FREIDENRICH
? STREET: 4365 EXECUTIVE DRIVE, SUITE 1600
? CITY: SAN DIEGO
? STATE: CALIFORNIA
? COUNTRY: USA
? ZIP: 92121
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/278,635B
? FILING DATE: 21-JUL-1994
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: REITER, STEPHEN E.
? REGISTRATION NUMBER: 31,192
? REFERENCE/DOCKET NUMBER: P41 9771
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619-677-1409
? TELEFAX: 619-677-1465
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 510 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-278-635B-4

```

```

Query Match      18.1%; Score 418.5; DB 1; Length 510;
Best Local Similarity 25.0%; Pred. No. 1.2e-35;
Matches 127; Conservative 88; Mismatches 178; Indels 115; Gaps 18;

QY 20 LQMLMGLMGLTSPGATATADPKNANVAKDLRLHAGLFTNYSDDVOPFGPTNVS L 79
DB 8 LQFWTHLYLMCLLP---AVLTQGGSHTHAEDRLFKHLFGYNNRMAPV-----PNTSD 59
QY 80 EMVVTY-----IDIDELNGKLTTHCMNLNRDEERWVQPSQYDNTQTTLKSSEVWT 132
DB 60 VVIVRFGLSIAQLIDVDKKNQMTTNWMLKQEMNDYVNRMDPAEFGVNTSLRVPSEMIWI 119
QY 133 PQTLLFNGDEGGLMAETQVTLSH---DGHFRMPRPVAVTYACELNMLNMPHDKOSCKLKI 189
DB 120 PDIVLYNNADGE-FAVYHMTKAHLFTGTVMVPRPAIYKSSCSIDYFFPFDQNCMKKF 178
QY 190 GSW---GLKYVLPNGARGESLDHDLVOSPEWEIYDSRAHFVSQDY-----YGMEX 240
DB 179 GSWTYDRAKIDLEO---MERTYDLKDYWESGEWAIINATGYNSKRYDCCAEIYDPVY 234
QY 241 TLTQRSSMTAVIYTPASIVILASAPFLPRHMGSEKIMINGLLITVIAALMFAQ 300
DB 235 YFVIRRLPLFTYINLLIPCLLSCLYVAVFLPSEC-GEKTLTLCISVLLSTVFLLLITE 293
QY 301 LPLVSNNTPLVAVFYESTSLYSVSTIVEVLYLATGKHK-----RLPEALRKLIHG 355
DB 294 IIFSTSLVPLIGEYLLFTMTFVTLSTIYIVFLNV---HHRSPSTHMMNWRVALLG 349
QY 356 HLGTLW---LSVSTGESQAEKT-----KEMDEHPYEADGE----- 392
DB 350 RVPMLAMNRPFLPMEHLGSPDLKLSPSYHLELTNMDAGEEETEEEBEEDENICVAGL 409
QY 393 -SSPLGINH-----TEVPGAKANQF----- 411

```

```

DB 410 PDSSMGVLYGHGLHRLAMEPETKP-SQASEILLSPOIKALEGVHYIADRLSEADS 468
QY 412 ----DMALLATAVDRISEFVSFLALII 435
DB 469 SVKEDMKYVNAVVDRI-----FLMLFIIV 492

RESULT 6
US-08-471-961-4
? Sequence 4, Application US/08471961
? Patent No. 6100046
? GENERAL INFORMATION:
? APPLICANT: ELGOYHEN, ANA BELEN
? APPLICANT: JOHNSON, DAVID S.
? APPLICANT: BOULTER, JAMES R.
? APPLICANT: HEINEMANN, STEPHEN F.
? TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL
? TITLE OF INVENTION: ACETYLCHOLINE-GATED ION CHANNEL RECEPTOR
? NUMBER OF SEQUENCES: 8
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: GRAY CARY WARE & FREIDENRICH
? STREET: 4365 EXECUTIVE DRIVE, SUITE 1600
? CITY: SAN DIEGO
? STATE: CALIFORNIA
? COUNTRY: USA
? ZIP: 92121
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/471,961
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/278,635
? FILING DATE: 21-JUL-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: REITER, STEPHEN E.
? REGISTRATION NUMBER: 31,192
? REFERENCE/DOCKET NUMBER: P41 9771
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619-677-1409
? TELEFAX: 619-677-1465
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 510 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-471-961-4

Query Match      18.1%; Score 418.5; DB 3; Length 510;
Best Local Similarity 25.0%; Pred. No. 1.2e-35;
Matches 127; Conservative 88; Mismatches 178; Indels 115; Gaps 18;

QY 20 LQMLMGLMGLTSPGATATADPKNANVAKDLRLHAGLFTNYSDDVOPFGPTNVS L 79
DB 8 LQFWTHLYLMCLLP---AVLTQGGSHTHAEDRLFKHLFGYNNRMAPV-----PNTSD 59
QY 80 EMVVTY-----IDIDELNGKLTTHCMNLNRDEERWVQPSQYDNTQTTLKSSEVWT 132
DB 60 VVIVRFGLSIAQLIDVDKKNQMTTNWMLKQEMNDYVNRMDPAEFGVNTSLRVPSEMIWI 119
QY 133 PQTLLFNGDEGGLMAETQVTLSH---DGHFRMPRPVAVTYACELNMLNMPHDKOSCKLKI 189
DB 120 PDIVLYNNADGE-FAVYHMTKAHLFTGTVMVPRPAIYKSSCSIDYFFPFDQNCMKKF 178
QY 190 GSW---GLKYVLPNGARGESLDHDLVOSPEWEIYDSRAHFVSQDY-----YGMEX 240
DB 179 GSWTYDRAKIDLEO---MERTYDLKDYWESGEWAIINATGYNSKRYDCCAEIYDPVY 234

```

Query Match	18.0%;	Score 416;	DB 1;	Length 529;
Best Local Similarity	27.6%;	Pred. No. 2.4e-35;		
Matches 109;	Conservative 75;	Mismatches 153;	Indels 58;	Gaps 12

```

1      RESULT      8
2      US-08-278-635B-3
3      : Sequence 3, Application US/08278635B
4      : Patent No. 5683912
5      :
6      : GENERAL INFORMATION:
7      :
8      : APPLICANT: ELGOYEN, ANA BELEN
9      : APPLICANT: JOHNSON, DAVID S.
10     : APPLICANT: BOULTER, JAMES R.
11     : APPLICANT: HEINEMANN, STEPHEN F.
12     :
13     : TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL
14     :
15     : TITLE OF INVENTION: ACETYLCHOLINE-GATED ION CHANNEL RECEPTOR
16     :
17     : NUMBER OF SEQUENCES: 8
18     :
19     : CORRESPONDENCE ADDRESS:
20     :
21     : ADDRESSEE: GRAY CARY WARE & FREIDENRICH
22     : STREET: 4365 EXECUTIVE DRIVE, SUITE 1600
23     : CITY: SAN DIEGO
24     : STATE: CALIFORNIA
25     :
26     : COUNTRY: USA
27     :
28     : ZIP: 92121
29     :
30     : COMPUTER READABLE FORM:
31     :
32     : MEDIUM TYPE: Floppy disk
33     : COMPUTER: IBM PC compatible
34     : OPERATING SYSTEM: PC-DOS/MS-DOS
35     : SOFTWARE: PatentIn Release #1.0, Version #1.25
36     :
37     : CURRENT APPLICATION DATA:
38     :
39     : APPLICATION NUMBER: US/08/278,635B
40     :
41     : FILING DATE: 21-JUL-1994
42     :
43     : CLASSIFICATION: 435
44     :
45     : ATTORNEY/AGENT INFORMATION:
46     :
47     : NAME: REITER, STEPHEN E.
48     : REGISTRATION NUMBER: 31,192
49     : REFERENCE/DOCKET NUMBER: P41 9771
50     :
51     : TELECOMMUNICATION INFORMATION:
52     :
53     : TELEPHONE: 619-677-1409
54     : TELEFAX: 619-677-1465
55     :
56     : INFORMATION FOR SEQ ID NO: 3:
57     :
58     : SEQUENCE CHARACTERISTICS:
59     :
60     : LENGTH: 457 amino acids
61     : TYPE: amino acid
62     :
63     : TOPOLOGY: linear
64     :
65     : MOLECULE TYPE: protein
66     :
67     : US-08-278-635B-3

```

Query Match 17.9%; Score 414.5; DB 1; Length 457;  
Best Local Similarity 25.8%; Pred. No. 2,7e-35;  
Matches 118; Conservative 97; Mismatches 170; Indels 73; Gaps 19;

QY 26 MLMLGLTSPGATATAPKNNANVAKDLRLAGLTNDSDVQPV---FQGTPTNVSLMV 82  
DB 7 LLLGLSSA-GLVIGSHEH-----RLVAKLFEDYSSVVRPEVDRHREIVQVTVGLQ- 57  
QY 83 VTYIDIDELNGKLTTHCMLNLRMRDEERWVQPSQYDNIQTITKSEVWTPQITLFGNDE 142  
DB 58 IQLINDEVNQIYTTNRLKQWVDYLNKNNPDYGVKKIHIPSEKIMRPDVLVYNAD 117  
QY 143 G--GLMAETOVTLSDGHFRMPRPAYTAYCELMLNMPDKOSCKLIKGSW---GLKV 196  
DB 118 GDEFAIVFTVLLDYTGHTITWTPPAIKSYCEIIVTHFPPEDEONCSMKLTGTWYDGSVA 177  
QY 197 VLPNGTARGESLDHDLNOSPMEIYDSRA--HFVSQDYG-----YME--YTLAQR 246  
DB 178 INPES-----DQPLSNFMSGEVWIKERGWKHV---FYSCPTPTTYDITVHFYMR 229  
QY 247 RSSWYTAIVTTPASCIYIALSAFWLPRHMGEEKIMINGLIIVIAFLMYFAQLLPVLS 306  
DB 230 LPLFYINVAIIPCLFSEFLSVLYLPTD--SGEKMTLSISVLSLYFLVLYELLPTS 288  
QY 307 NMPPLVIVYSTLSLYSVSTIVEVLYLATGKHRR-----LPEALRLHLGHGTWL 361  
DB 289 SAVPLIGKMYLFTWVFVIAIIITVIVI---NTHHRSPTTHMPEVVRKVFIDITPINM 344  
QY 362 LLSVFTTGSQAEK---TDEMDEHPYEADDEOSSPLG-----INTEVPGA----- 406  
DB 345 FFSMTKRPSRDKOKRIETDIDIS--DISGKPGPPMGHSPLIKHPKSAIEGVKYI 402  
QY 407 -----KANQ-----FDWALLATAVDRISEVFSFLAFLI 434  
DB 403 AETMKSDQESNNAEMKYVAMWMDHLLGVFMVLCI 440

RESULT 9  
US-08-471-961-3  
; Sequence 3, Application US/08471961  
; Patent No. 6100046  
; GENERAL INFORMATION:  
; APPLICANT: ELGOYHEN, ANA BELEN  
; APPLICANT: JOHNSON, DAVID S.  
; APPLICANT: BOULTER, JAMES R.  
; APPLICANT: HEINEMANN, STEPHEN F.  
; TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GRAY CARY WARE & FREIDENRICH  
; STREET: 4365 EXECUTIVE DRIVE, SUITE 1600  
; CITY: SAN DIEGO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 92121  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471,961  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,635  
; FILING DATE: 21-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: REITER, STEPHEN E.  
; REGISTRATION NUMBER: 31,192  
; REFERENCE/DOCKET NUMBER: P41 9771

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-677-1409  
; TELEFAX: 619-677-1465  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 457 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-471-961-3

Query Match 17.9%; Score 414.5; DB 3; Length 457;  
Best Local Similarity 25.8%; Pred. No. 2,7e-35;  
Matches 118; Conservative 97; Mismatches 170; Indels 73; Gaps 19;

QY 26 MLMLGLTSPGATATAPKNNANVAKDLRLAGLTNDSDVQPV---FQGTPTNVSLMV 82  
DB 7 LLLGLSSA-GLVIGSHEH-----RLVAKLFEDYSSVVRPEVDRHREIVQVTVGLQ- 57  
QY 83 VTYIDIDELNGKLTTHCMLNLRMRDEERWVQPSQYDNIQTITKSEVWTPQITLFGNDE 142  
DB 58 IQLINDEVNQIYTTNRLKQWVDYLNKNNPDYGVKKIHIPSEKIMRPDVLVYNAD 117  
QY 143 G--GLMAETOVTLSDGHFRMPRPAYTAYCELMLNMPDKOSCKLIKGSW---GLKV 196  
DB 118 GDEFAIVFTVLLDYTGHTITWTPPAIKSYCEIIVTHFPPEDEONCSMKLTGTWYDGSVA 177  
QY 197 VLPNGTARGESLDHDLNOSPMEIYDSRA--HFVSQDYG-----YME--YTLAQR 246  
DB 178 INPES-----DQPLSNFMSGEVWIKERGWKHV---FYSCPTPTTYDITVHFYMR 229  
QY 247 RSSWYTAIVTTPASCIYIALSAFWLPRHMGEEKIMINGLIIVIAFLMYFAQLLPVLS 306  
DB 230 LPLFYINVAIIPCLFSEFLSVLYLPTD--SGEKMTLSISVLSLYFLVLYELLPTS 288  
QY 307 NMPPLVIVYSTLSLYSVSTIVEVLYLATGKHRR-----LPEALRLHLGHGTWL 361  
DB 289 SAVPLIGKMYLFTWVFVIAIIITVIVI---NTHHRSPTTHMPEVVRKVFIDITPINM 344  
QY 362 LLSVFTTGSQAEK---TDEMDEHPYEADDEOSSPLG-----INTEVPGA----- 406  
DB 345 FFSMTKRPSRDKOKRIETDIDIS--DISGKPGPPMGHSPLIKHPKSAIEGVKYI 402  
QY 407 -----KANQ-----FDWALLATAVDRISEVFSFLAFLI 434  
DB 403 AETMKSDQESNNAEMKYVAMWMDHLLGVFMVLCI 440

RESULT 10  
US-08-464-258B-3  
; Sequence 3, Application US/08464258B  
; Patent No. 6013766  
; GENERAL INFORMATION:  
; APPLICANT: ELGOYHEN, ANA BELEN  
; APPLICANT: JOHNSON, DAVID S.  
; APPLICANT: BOULTER, JAMES R.  
; APPLICANT: HEINEMANN, STEPHEN F.  
; TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GRAY CARY WARE & FREIDENRICH LLP  
; STREET: 4365 EXECUTIVE DRIVE, SUITE 1600  
; CITY: SAN DIEGO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 92121  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25



CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/464,258B  
 FILING DATE: 06/05/95  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/278,635  
 FILING DATE: 21-JUL-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: REITER, STEPHEN E.  
 REGISTRATION NUMBER: 31,192  
 REFERENCE/DOCKET NUMBER: P41 9989  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 619-677-1409  
 TELEFAX: 619-677-1465  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 458 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-464-258B-3

Query Match 17.9%; Score 414; DB 3; Length 458;  
 Best Local Similarity 25.7%; Pred. No. 3e-35;  
 Matches 118; Conservative 97; Mismatches 170; Indels 74; Gaps 19;

QY 26 MLMLGTSVPGATATADKRNANVAKLDRHAGLFNTYDSQVY---FQGTPTNWSLEAV 82  
 DB 7 LLLGLISSA-GVILGSEHET-----RLVAKLFEDYSVVRPVDHREIVQVTVGLD- 57  
 QY 83 VTYIDIDELNGKLTTHCHMLNLRMDEERYWQSOYDNTQITLKSEVTPQITFENGDE 142  
 DB 58 IOLINDEVNOQVITTNVRLKQOWVYNLKNPDDYGGVKKIIPSEKIRPVPVLYXNAD 117  
 QY 143 G--GMAETQVTLSDGHFRMPRPVATYACELNMLNMPHDQSKLKIGSW---GLKV 196  
 DB 118 GFALVKKFKVLDDYTGHTWTPRAIFKSCYCEIIVTHFPFDQNSMKLGTYDGSVVA 177  
 QY 197 VLPENGTAGESLDHDDLVOSEPELVDRA--HFVSQDYG-----YME--YTLTAOR 246  
 DB 178 INPE-----DQPDLSNEMSEGENVKEKRGKHNW---FYSCPTPYLDITVHFVMQR 229  
 QY 247 RSSMTAVIYTPASCIVILALSAFWLPRMGEKIMINGLLIIVIAFLMYAQLLPVLS 306  
 DB 230 LPLRYIVANVITPCLFSTLSVFTLPD-SEKMTLSISVLSLTVFLVVELIPSTIS 288  
 QY 307 NNTPLVIVFYSTSLYLSVSTIVEVLYLATGKHRR---LPEALKRLHGHGLGWL 361  
 DB 289 SAVPLIGKMYLFTWVFVLAIIIVYI---NTHHRSPSTHIMPEWAKVFDITIPNIM 344  
 QY 362 LLSVSTGESQAEK---TKMDEHPYEADEQESSPLG-----INTEVGA----- 406  
 DB 345 FFSYTKRPSRDKQEKRIFFEDIS--DISGKPGPPPMGFHSPLIKHPVSAIEGVKY 402  
 QY 407 -----KANO-----EDWALLATAVDRISEFSLAFLI 434  
 DB 403 IAEYKSDQESNNAABEKYVAMVMDHILGVFMLVCLT 441

RESULT 11  
 US-08-466-589-2  
 ; Sequence 2, Application US/08466589  
 ; Patent No. 5837489  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Elliot, Kathryn J.  
 ; APPLICANT: Ellis, Steven B.  
 ; APPLICANT: Harpold, Michael M.  
 ; TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Brown, Martin, Haller & McClaim

STREET: 1660 Union Street  
 CITY: San Diego  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 92101-2926  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/466,589  
 FILING DATE: June 5, 1995  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/028,031  
 FILING DATE: March 8, 1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Seidman, Stephanie L.  
 REGISTRATION NUMBER: 33,779  
 REFERENCE/DOCKET NUMBER: 6362-9950  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 619-238-0999  
 TELEFAX: 619-238-0062  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 528 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 US-08-466-589-2

Query Match 17.8%; Score 412.5; DB 2; Length 528;  
 Best Local Similarity 27.4%; Pred. No. 5.5e-35;  
 Matches 108; Conservative 76; Mismatches 153; Indels 57; Gaps 12;

QY 13 GGPLLLLOML-MGLMLGTSVPGATATADPKNA-----NKKALDR 53  
 DB 2 GSCCVFLSFTKLSIMWLLITPAGGEAKRPPRAPGDLSPSPALAPQGSHTETEDR 61  
 QY 54 LHAGFTNTDSDVOYVFGSTPTNWSLEAVYV-----IDDELNGKLTTHCHMLNLR 106  
 DB 62 LFKHLFRGYNRRARVY-----PNTSDVIVRFGLSIAQLIDVDKNOQMTTVMWLKQES 116  
 QY 107 DEERYWQSOYDNTQITLKSEVTPQITFENGDEGGMATQVTLSH---DGHFRMP 163  
 DB 117 DYKLRMPADPFCNITSLRPSSEMIMPDIVLYN--KXGEFAVTHMTKALHFGTGVHWP 174  
 QY 164 PAVYATYCELNMLNMPHDQSKLKIGSW--GLKVLPENGTAGESLDHDDLVOSEPBW 220  
 DB 175 PAIVKSSGSDVTFPPFDQNCMKFGSWTYDKARIDLEQ---MEQFDLDKYWESGHW 230  
 QY 221 EIVDSRAHVQSDY-----YGYMEYTLAQRSSMTAVIYTPASCIVILALSAFWL 274  
 DB 231 AIVNATGYNSKKKYDCCAEIYPDVYAFVIRLPLFTIINLLIIPCLLSCLTVLVFLPS 290  
 QY 275 HMGGEKIMINGLLIIVIAFLMYFAQLPVLSSNNTPLVIVFYSTSLYLSVSTIVEVLYL 334  
 DB 291 DC-GEKITLCISVLSLTVFELLITIEIIPSTSLVYIPLIGEYLLFMIFVTLISVITFVL 349  
 QY 335 YLATGKKH---RLPEALKRLHGHGLGWL 364  
 DB 350 NV---DHRSPSTHTMPHWVKGALLGCVPRMLMN 380

RESULT 12  
 US-08-700-636-2  
 ; Sequence 2, Application US/08700636  
 ; Patent No. 5910582  
 ; GENERAL INFORMATION:

```

? APPLICANT: Elliot, Kathryn J.
? APPLICANT: Ellis, Steven B.
? APPLICANT: Harpold, Michael M.
? TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
? NUMBER OF SEQUENCES: 12
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
? STREET: 444 South Flower Street, Suite 2000
? CITY: Los Angeles
? STATE: CA
? COUNTRY: USA
? ZIP: 90071
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/700,636
? FILING DATE: 16-JUL-1996
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/028,031
? FILING DATE: 08-MAR-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Reiter, Stephen E.
? REGISTRATION NUMBER: 31,192
? REFERENCE/DOCKET NUMBER: P41 9368
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619-546-4737
? TELEFAX: 619-546-9392
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 528 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: protein
? US-08-700-636-2

```

```

Query Match      17.8%; Score 412.5; DB 2; Length 528;
Best Local Similarity 27.4%; Pred. No. 5,5e-35;
Matches 108; Conservative 76; Mismatches 153; Indels 57; Gaps 12;

QY 13 GPGPLPILQL-MGMLMGITSVGATATADPKNA-----NVRALDR 53
b 2 GPCSPVLFSTFKLSIMWLLITPAGGEAKRPPRAPGDPSPSPALPOGGSHTETEDR 61
QY 54 LHAGLFTNYSDVQVPVQGTPTNVSLFEMVTV-----IDDELNGKLTTHQWMLNRM 106
Db 62 LFNHLFRGYRMARPV-----PNTSDVIVIRFGISIAQLIDVDEKQNMNTNWLKQEW 116
QY 107 DEERVWQPSQYDNIQTOTLKSSSEWTPQITLFGNDEGLMAETQVLSH---DGHFRMP 163
Db 117 DYLRNMPADFGNITSLRVSEMIWIPDIVLYN--KXGEFAVTHMTRKHLFSTGTVHWVP 174
QY 164 PAVYTAVCENLMLNMPHDKOSKLIKGSW---GLKVVLPENGTARGESLDHDDLVOGSP 220
Db 175 PAIYKSSCIDVYFFPDQNCMKKFGSWYDKAKIDLEO---MEOYVLDKDYWESGEW 230
QY 221 EIVDSRAHFVSODY-----YGYMEYTLTAQRSSMYTAVIYTPASCIIVLALSAFWLPP 274
Db 231 AIYNATGYNSKKYDCCAEIYPDVTAFAVIRLPLEFTYINLIPCLLISCLTVLVEFLPS 290
QY 275 HMGCEKIMINGLLIYVIAALMFPAQLPVLNSNTPVLVIFYSTSLIYLSVSTIVEVYL 334
Db 291 DC-GEKTLICISVLSLTVFLLEITELIPSTSVIPIGCEYLLFTMTFVLSIVTVEVL 349
QY 335 YLATGKHK---RRLPEALRKLHGHGLGFWLLS 364
Db 350 NV---DHRSPSTHTMPHWVIRGALIGCVPRWLLMN 380

```

```

RESULT 13
US-08-467-574-2
? Sequence 2, Application US/08467574
? Patent No. 6022704
? GENERAL INFORMATION:
? APPLICANT: Elliot, Kathryn J.
? APPLICANT: Ellis, Steven B.
? APPLICANT: Harpold, Michael M.
? TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
? NUMBER OF SEQUENCES: 12
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Brown, Martin, Haller & McClaim
? STREET: 1660 Union Street
? CITY: San Diego
? STATE: CA
? COUNTRY: USA
? ZIP: 92101-2926
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? OPERATING SYSTEM: DOS
? SOFTWARE: FastSeq Version 1.5
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/467,574
? FILING DATE: June 5, 1995
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/028,031
? FILING DATE: March 8, 1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Seidman, Stephanie L.
? REGISTRATION NUMBER: 33,779
? REFERENCE/DOCKET NUMBER: 6362-9949
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619-238-0999
? TELEFAX: 619-238-0062
? TELEX:
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 528 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: protein
? US-08-467-574-2

```

```

Query Match      17.8%; Score 412.5; DB 3; Length 528;
Best Local Similarity 27.4%; Pred. No. 5,5e-35;
Matches 108; Conservative 76; Mismatches 153; Indels 57; Gaps 12;

QY 13 GPGPLPILQL-MGMLMGITSVGATATADPKNA-----NVRALDR 53
b 2 GPCSPVLFSTFKLSIMWLLITPAGGEAKRPPRAPGDPSPSPALPOGGSHTETEDR 61
QY 54 LHAGLFTNYSDVQVPVQGTPTNVSLFEMVTV-----IDDELNGKLTTHQWMLNRM 106
Db 62 LFNHLFRGYRMARPV-----PNTSDVIVIRFGISIAQLIDVDEKQNMNTNWLKQEW 116
QY 107 DEERVWQPSQYDNIQTOTLKSSSEWTPQITLFGNDEGLMAETQVLSH---DGHFRMP 163
Db 117 DYLRNMPADFGNITSLRVSEMIWIPDIVLYN--KXGEFAVTHMTRKHLFSTGTVHWVP 174
QY 164 PAVYTAVCENLMLNMPHDKOSKLIKGSW---GLKVVLPENGTARGESLDHDDLVOGSP 220
Db 175 PAIYKSSCIDVYFFPDQNCMKKFGSWYDKAKIDLEO---MEOYVLDKDYWESGEW 230
QY 221 EIVDSRAHFVSODY-----YGYMEYTLTAQRSSMYTAVIYTPASCIIVLALSAFWLPP 274
Db 231 AIYNATGYNSKKYDCCAEIYPDVTAFAVIRLPLEFTYINLIPCLLISCLTVLVEFLPS 290

```

```

0Y      54 LHAGLETFNYSDSDVOPVFGGPTNYSLEWVY-----IDIDELNGKLTTHCVLNRWR 106
      |  || | : :|| | | :|| | ||: || | |

```

44 KNAVKALDRHAGLF7TNYSDVQPFQGTPT--NVSLEWVYI-DIDELNGKLTTHCN 100  
: || | : | : | : || | : | : | : | : | : | : | : | : | : |



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 08:50:47 ; Search time 79.64 Seconds  
(without alignments)  
615.062 Million cell updates/sec

Title: US-09-732-680A-2  
Perfect score: 2311  
Sequence: 1 MTTTPRIKAPVSGPLILL.....RISFVSFLAILAIRCSV 441

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
A.Geneseq\_032802:\*  
1: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/genesep-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/genesep-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2311	100.0	441	22	ABB68475
2	2311	100.0	441	22	ABB68218
3	420	18.2	627	18	AAW11825
4	418	18.1	627	18	AAW11824
5	416	18.0	529	15	AAW09023
6	415	18.0	529	16	AAW44155
7	416	18.0	529	18	AAW73966
8	415	18.0	529	18	AAW09021
9	412	17.8	450	22	AA67161
10	409	17.7	450	22	AAU00403
11	409	17.7	450	22	AAE00280

12	405.5	17.5	457	6	AAE50361
13	405.5	17.5	457	21	AAE56386
14	404.5	17.5	433	22	AAU00402
15	402.5	17.4	519	22	ABE62694
16	399.5	17.3	437	22	AAU69151
17	398.5	17.2	468	18	AAW09024
18	397.5	17.2	502	22	AAE51021
19	397	17.2	479	17	AAE89358
20	395.5	17.1	457	22	AAU00404
21	394	17.0	494	18	AAW09018
22	392.5	17.0	436	11	AAE06236
23	391	16.9	458	18	AAW09020
24	391	16.9	458	22	AAE12775
25	391	16.9	538	22	ABG27365
26	391	16.9	627	15	AAW44152
27	383.5	16.6	501	22	AAW78851
28	383.5	16.6	506	22	ABE11811
29	383.5	16.6	506	22	AAW79835
30	383	16.6	456	11	AAE06255
31	381	16.5	456	11	AAE06254
32	378.5	16.4	461	5	AAE40417
33	378	16.4	479	22	AAE12823
34	377.5	16.3	449	17	AAE86421
35	375.5	16.2	449	14	AAE35057
36	374	16.2	502	22	AAE50016
37	373	16.1	502	15	AAW44153
38	373	16.1	502	18	AAW09025
39	373	16.1	502	19	AAE69216
40	373	16.1	502	21	AAE24088
41	373	16.1	502	22	AAE82690
42	372	16.1	502	22	AAE50012
43	372	16.1	502	22	AAE50017
44	371	16.1	502	22	AAE50015
45	370.5	16.0	504	18	AAW09022

## ALIGNMENTS

RESULT 1		
ABBB68475	ID	ABBB68475 standard; Protein: 441 AA.
XX	XX	
AC	ABBB68475;	
XX	XX	
DT	26-MAR-2002 (first entry)	
XX	XX	
DE	Drosophila melanogaster polypeptide SEQ ID NO 32217.	
XX	XX	
KW	Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.	
XX	XX	
CS	Drosophila melanogaster.	
XX	XX	
PN	WO200171042-A2.	
XX	XX	
PD	27-SEP-2001.	
XX	XX	
PF	23-MAR-2001; 2001WO-US09231.	
XX	XX	
PR	23-MAR-2000; 2000US-191637P.	
PR	11-JUL-2000; 2000US-0614150.	
XX	XX	
PA	(PEKE ) PE CORP NY.	
XX	XX	
PI	Venter JC, Adams M, Li PMD, Myers EW;	
XX	XX	
WPI	2001-656860/75.	
DR	N-PSDB; ABL12578.	
XX	XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -	

Human acetyl choli  
Human acetylcholin  
Mature human alpha  
Drosophila melanog  
Human Acetylcholin  
Neuronal nicotinic  
JTF-38 nAChR alpha  
Alpha-9 nicotinic  
Human alpha nicoti  
Neuronal nicotinic  
Calf acetylcholine  
Neuronal nicotinic  
Human cholinergic  
Novel human diagno  
Human neuronal nic  
Human protein SEQ  
Human acetylcholin  
Human protein SEQ  
Mouse acetylcholin  
Human acetylcholin  
Acetyl choline acc  
Caenorhabditis ele  
Acetylcholine rece  
Alpha subunit of T  
Mutant human alpha  
Human neuronal nic  
Neuronal nicotinic  
V274T variant huma  
Human PRO2145 prot  
Nicotinic acetylch  
Wild-type human al  
Mutant human alpha  
Mutant human alpha  
Neuronal nicotinic

XX Disclosure; SEQ ID NO 32217; 21pp + Sequence Listing; English.  
PS  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB101840-AB101845), expressed DNA  
CC (AB057737-AB057742).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
XX  
SQ Sequence 441 AA:  
  
Query Match 100.0%; Score 2311; DB 22; Length 441;  
Best Local Similarity 100.0%; Pred. No. 5.1e-227;  
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MTTTPKIKAPSGPGLLLQMLMGLMGLTSPGATATADPKNAVKAIDRLHGLFT 60  
Db 1 mtttppkikapsgrpjlllqmlmgllmgltsvpgatacadpknanvkaidrlhaglft 60  
QY 61 NYDSVQPVFGPPTNVSLEKVVYIDIDELNGKLTTHCWLNRDEERYWQPSQYDNI 120  
Db 61 nydsdvpgpvtgnvslekvvyididelngrltthcwlntwdeerywqpsqydni 120  
QY 121 TQTLKSEVWTPQITLFGNDEGGLMAETQYTLSDGHFRMPRAVYATACELMMLMPH 180  
Db 121 tqtlksewvtpqitlfnngdegglmaetqytlshdghfrmppravyataycelmlmnp 180  
QY 181 DKOSCKIKISGKAVLPENGTARGESLDHDLVQSPMEIYDSRAHFVSQDYGYMEY 240  
Db 181 dkgsckikisgkavlpengtaragesldhdvqspeweiysdranhfvsqdygymey 240  
QY 241 TLTAQRSSMYTAVIYTPASCIIVLALSAFWLPPHMGGEKIMINGLLIIVIAFLMYAQ 300  
Db 241 tltagrssmytavlytpasciivlalsafwlpphmggekimgllliviiaeflmyaq 300  
QY 301 LLEPLSNNTPLVYIFSTSLIYSVTIVEVLYLATGKHRRLPALRKLHGLGTM 360  
Db 301 lleplsnntplvlyfstsllysvtivevlylatgkhrrlpalkrklhghlgtw 360  
QY 361 LLLSVFTGSGQAEKTKEMDEHPYEADQESSPLGINTEVPGAKANOFDMALLATAV 420  
Db 361 lllsvftsgsqaektkemdenhyeadeqessplginntevpgakangdwallatav 420  
QY 421 DRISFVSFLAFLILAIRCSV 441  
Db 421 drisfvsflaflilaircsv 441  
  
RESULT 2  
ID AAB86218 standard; Protein: 441 AA.  
XX  
AC AAB86218:  
XX  
DT 31-AUG-2001 (first entry)  
XX  
DE D. melanogaster acetylcholine receptor beta subunit Db3 protein.  
XX  
KM Acetylcholine receptor beta subunit; Db3; fruitfly; insecticidal;  
XX transgenic; modulator; plant-protection agent; medicine.  
XX  
OS *Drosophila melanogaster*.  
XX  
PN EP1106689-A2.  
XX  
PD 13-JUN-2001.

XX  
XX 28-NOV-2000; 2000EP-0125300.  
XX  
XX 10-DEC-1999; 99DE-1059582.  
XX  
XX (FARB) BAYER AG.  
XX  
PI Adamczewski M, Methfessel C, Schulte T;  
XX WPI: 2001-376906/40.  
XX DR N-PSDB; AAH20992.  
XX  
XX New nucleic acid encoding *Drosophila* acetylcholine receptor subunit,  
XX useful in screening for plant protection agents and pharmaceuticals  
XX  
XX Claim 9; Page 13-15; 18pp; German.  
XX  
XX This invention describes a novel nucleic acid (I) which encodes a  
XX *Drosophila* melanogaster acetylcholine receptor beta subunit, Db3 which  
XX has insecticidal activity. (I), related host cells, polypeptides, ACR,  
XX antibodies, transgenic invertebrates and regulatory regions, are used to  
XX identify agents that modulate ACR activity, potentially useful as  
XX plant-protection agents (e.g. insecticides) or pharmaceuticals for human  
XX or veterinary medicine. This sequence represents the *Drosophila*  
XX melanogaster acetylcholine receptor beta subunit described in the method  
XX of the invention.  
XX  
SQ Sequence 441 AA:  
  
Query Match 100.0%; Score 2311; DB 22; Length 441;  
Best Local Similarity 100.0%; Pred. No. 5.1e-227;  
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MTTTPKIKAPSGPGLLLQMLMGLMGLTSPGATATADPKNAVKAIDRLHGLFT 60  
Db 1 mtttppkikapsgrpjlllqmlmgllmgltsvpgatacadpknanvkaidrlhaglft 60  
QY 61 NYDSVQPVFGPPTNVSLEKVVYIDIDELNGKLTTHCWLNRDEERYWQPSQYDNI 120  
Db 61 nydsdvpgpvtgnvslekvvyididelngrltthcwlntwdeerywqpsqydni 120  
QY 121 TQTLKSEVWTPQITLFGNDEGGLMAETQYTLSDGHFRMPRAVYATACELMMLMPH 180  
Db 121 tqtlksewvtpqitlfnngdegglmaetqytlshdghfrmppravyataycelmlmnp 180  
QY 181 DKOSCKIKISGKAVLPENGTARGESLDHDLVQSPMEIYDSRAHFVSQDYGYMEY 240  
Db 181 dkgsckikisgkavlpengtaragesldhdvqspeweiysdranhfvsqdygymey 240  
QY 241 TLTAQRSSMYTAVIYTPASCIIVLALSAFWLPPHMGGEKIMINGLLIIVIAFLMYAQ 300  
Db 241 tltagrssmytavlytpasciivlalsafwlpphmggekimgllliviiaeflmyaq 300  
QY 301 LLEPLSNNTPLVYIFSTSLIYSVTIVEVLYLATGKHRRLPALRKLHGLGTM 360  
Db 301 lleplsnntplvlyfstsllysvtivevlylatgkhrrlpalkrklhghlgtw 360  
QY 361 LLLSVFTGSGQAEKTKEMDEHPYEADQESSPLGINTEVPGAKANOFDMALLATAV 420  
Db 361 lllsvftsgsqaektkemdenhyeadeqessplginntevpgakangdwallatav 420  
QY 421 DRISFVSFLAFLILAIRCSV 441  
Db 421 drisfvsflaflilaircsv 441  
  
RESULT 3  
ID AAW11825 standard; Protein: 627 AA.  
XX  
AC AAW11825:  
XX  
PD AAW11825.

```

DT 05-OCT-1997 (first entry)
XX Alpha4 subunit of mutant nAChR.
XX nAChR; mutation; autosomal dominant nocturnal frontal lobe epilepsy;
XX ADNFLE; neuronal acetylcholine receptor.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 280 /note= "Ser 248 Phe is the ADNFLE-
XX inducing mutation (see AAM11824)"
XX
XX A09656247-A.
XX
XX 09-JAN-1997.
XX
XX 28-JUN-1996; 96AU-0056247.
XX
XX 28-JUN-1995; 95AU-0003840.
XX
XX (UYBO-) UNIV BONN.
XX (UYME-) UNIV MELBOURNE.
XX (WOME-) WOMEN'S & CHILDREN'S HOSPITAL.
XX
XX Berkovic SF, Mulley JC, Phillips HA, Propping PJ;
XX Scheffer IE, Steinleinok, Sutherland GR, Wallace RH;
XX
XX WPI: 1997-100506/10.
XX
XX N-PSDB; AAT59528.
XX
XX CHRNA4 gene encodes neuronal nicotinic acetylcholine receptor
XX fragment - used in diagnosis of autosomal dominant nocturnal frontal
XX lobe epilepsy
XX
XX Disclosure: Fig 13; 20pp; English.
XX
XX The DNA sequence of the normal nAChR is given in AAT59527,
XX the DNA sequence of the ADNFLE-inducing mutant is given in AAT59528.
XX Mutations at codon 248, which have the effect of replacing
XX serine by another amino acid (e.g. phenylalanine) in the sixth
XX amino acid of the transmembrane domain 2 (M2) of the alpha4
XX subunit of nAChR have been found to be associated with ADNFLE.
XX The primers given in AAT59529 and AAT59530 were used in the
XX amplification of part of exon 5.
XX
XX Sequence 627 AA:
SQ

```

Query Match 18.2%; Score 420; DB 18; Length 627;  
 Best Local Similarity 29.8%; Pred. No. 1.2e-33;  
 Matches 102; Conservative 67; Mismatches 139; Indels 34; Gaps 9;

```

OY 11 VSGPGLPILLQMLMGMLGLTSVPGATATADPKNANVAKALRLHAGLFTNTDSDVPYF 70
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 3 1ggpgaprlpplllllgtgllraashvet-----rahaerlllklifsgynksrpy- 55
OY 71 OGPTPNVSLSEMYVY-----IDIDELNGKLTTTCWMLRNRDEBRWQPSQOYDNTIOI 123
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 56 ----anisdvnlvtrfglsiaqlidvdeknqmtlnwvkgewhdyklrvdpadyentstl 111
OY 124 TLKSSRWTPQITLTFNGDGGGLMAE--QVYTLSDGCHFRMMPAVYTATCELMNLMNPHD 181
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 112 rlpseilwppdlivynnaagdfavthltkahlfhdgivgwcpptalyksscsldvltfptd 171
OY 182 KOSCKLIKGSW--GLKVVLPENGTARGESLDHDLVQSPEREIVDSRAHFVSODY----- 234
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 172 gqncemkfgswtydkakdl-vnmhar---vqldlwegevwivdavgtyntirkyecca 227
OY 235 --GYMEVYTLAQRSSMTAVIYTPASCIVITLALSAFLLPRMGGEKTMINGLLIIVIA 292
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 228 e1ypdltyafvtrfplrfyctlnllipclliscvlvfy1psec-gekltlclfvllslit 286

```

```

OY 293 AFLMEFAQLPLVLSNNPLVWIFYSTLSLYLSTVSTIVELVL 334
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 287 vlllltelipstslvpligeyllfmlfvslsvlcvfvl 328

```

RESULT 4  
 AAM11824  
 ID AAM11824 standard; Protein: 627 AA.  
 XX  
 XX AAM11824;  
 AC  
 AC AAM11824;  
 DT 05-OCT-1997 (first entry)  
 XX  
 XX Alpha4 subunit of normal nAChR.  
 XX  
 XX nAChR; mutation; autosomal dominant nocturnal frontal lobe epilepsy;  
 XX ADNFLE; neuronal acetylcholine receptor.  
 XX  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX Misc-difference 280 /note= "Ser 248 Phe is the ADNFLE-  
 XX inducing mutation (see AAM11825)"  
 XX  
 XX A09656247-A.  
 XX  
 XX 09-JAN-1997.  
 XX  
 XX 28-JUN-1996; 96AU-0056247.  
 XX  
 XX 28-JUN-1995; 95AU-0003840.  
 XX  
 XX (UYBO-) UNIV BONN.  
 XX (UYME-) UNIV MELBOURNE.  
 XX (WOME-) WOMEN'S & CHILDREN'S HOSPITAL.  
 XX  
 XX Berkovic SF, Mulley JC, Phillips HA, Propping PJ;  
 XX Scheffer IE, Steinleinok, Sutherland GR, Wallace RH;  
 XX  
 XX WPI: 1997-100506/10.  
 XX  
 XX N-PSDB; AAT59527.  
 XX  
 XX CHRNA4 gene encodes neuronal nicotinic acetylcholine receptor  
 XX fragment - used in diagnosis of autosomal dominant nocturnal frontal  
 XX lobe epilepsy  
 XX  
 XX Disclosure: Fig 13; 20pp; English.  
 XX  
 XX The DNA sequence of the normal nAChR is given in AAT59527,  
 XX the DNA sequence of the ADNFLE-inducing mutant is given in AAT59528.  
 XX Mutations at codon 248, which have the effect of replacing  
 XX serine by another amino acid (e.g. phenylalanine) in the sixth  
 XX amino acid of the transmembrane domain 2 (M2) of the alpha4  
 XX subunit of nAChR have been found to be associated with ADNFLE.  
 XX The primers given in AAT59529 and AAT59530 were used in the  
 XX amplification of part of exon 5.  
 XX  
 XX Sequence 627 AA:  
 SQ

Query Match 18.1%; Score 418; DB 18; Length 627;  
 Best Local Similarity 29.8%; Pred. No. 1.9e-33;  
 Matches 102; Conservative 67; Mismatches 139; Indels 34; Gaps 9;

```

OY 11 VSGPGLPILLQMLMGMLGLTSVPGATATADPKNANVAKALRLHAGLFTNTDSDVPYF 70
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 3 1ggpgaprlpplllllgtgllraashvet-----rahaerlllklifsgynksrpy- 55
OY 71 OGPTPNVSLSEMYVY-----IDIDELNGKLTTTCWMLRNRDEBRWQPSQOYDNTIOI 123
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 56 ----anisdvnlvtrfglsiaqlidvdeknqmtlnwvkgewhdyklrvdpadyentstl 111

```





XX PS Disclosure; Page 68-70; 99pp; English.

XX CC The present sequence represents a human neuronal nicotinic acetylcholine

XX CC receptor (nAChR) subunit. The cells expressing the alpha and/or beta

XX CC nAChR subunits may be used in a method of screening compounds to

XX CC identify any which modulate the activity of human neuronal nAChR.

XX CC Subunit specific antibodies may be used to monitor the distribution

XX CC and expression density of various subunits in normal vs diseased brain

XX CC tissues. Testing of single receptor subunits or specific receptor

XX CC subunit combinations with a variety of potential agonists or antagonists

XX CC provides information with respect to the function and activity of the

XX CC individual subunits and should lead to the identification and design of

XX CC compounds that are capable of very specific interaction with one or

XX CC more receptor subtypes. The resulting drugs should exhibit fewer

XX CC unwanted side effects than drugs identified e.g. screening with cells

XX CC that express a variety of subtypes.

XX SQ Sequence 529 AA;

Query Match 18.0%; Score 416; DB 15; Length 529;  
 Best Local Similarity 27.6%; Pred. No. 2.3e-33;  
 Matches 109; Conservative 75; Mismatches 153; Indels 58; Gaps 12;

QY 13 GPGCLPILQLML-MGMLMGITSVPGATATADPKNA-----NKKALDR 53  
 DB 2 GPSCPVILSFKLISLWLLIPAGGEEKRPPTAPGDLSPSPALPGGSHTEDE 61  
 QY 54 LHAGLFTNYDSVQVFPQGTPTNVSLKENVVY-----IDIDELNGKLTTHCWLNLNR 106  
 DB 62 Lfkhlfgrynwarpv-----pntsdvivrfglsiaqlldvdeknqmttnvwlkqgs 116  
 QY 107 DEERWQPSQVDNITQITLKSEVWTPQITLFGNGDEGGLMAETQVTLSH---DGHFRMP 163  
 DB 117 dyklirmpadfgnltslvpsenmipdlvlynadg-favltmktkahlstgtvhwvp 175  
 QY 164 PAVTAYCELMLNMPHDKOSCKLRIGSM---GLKVLPENGTARGESLDHDLVQSPDW 220  
 DB 176 palysscsidvtfpfqgnckmkfgswtydkakidleg---meqtdldkdyesgew 231  
 QY 221 EIVDSRAHFVSODY-----YGYMEYTLTAQRSSMYTAVIYTPASCIIVIALSAFWLP 274  
 DB 222 alvatgynskkydcaaeipdvcyafvtrrplpfyctlnllipclliscslvlfyips 291  
 QY 275 HMGGEKIMINGLLIIVIAFLMFAQLLPVLSNNTPLVVIFYSTSLVSTIVEVLV 334  
 DB 292 dc-geklclclslvllstlvtlllletlpsstslvpllgeyillfemifvltslvltvyl 350  
 QY 335 YLATGKHKR-----RLPEALRKLHGHGCTWLLS 364  
 DB 351 nv-----hnrspsthmphwvrgallgcprwllm 381

RESULT 7

AA073966 ID AAR73966 standard; Protein: 529 AA.

XX AC AAR73966;

XX DT 30-NOV-1995 (first entry)

XX DE Alpha 2 subunit of a human neuronal nicotinic acetylcholine receptor.

XX KW Human nAChR; neuronal nicotinic acetylcholine receptor;

XX RM neurotransmitter.

XX OS Homo sapiens.

XX PN W09513299-A.

XX PD 18-MAY-1995.

PF 08-NOV-1994; 94WO-US12859.

XX RR 08-NOV-1993; 93US-0149503.

XX PA (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.

XX PI Elliott RJ, Ellis SB, Harpold MW;

XX DR WPI; 1995-194036/25.

XX N-PSDB; AA090387.

XX PT New human neuronal nicotinic acetylcholine receptor alpha2 subunit DNA  
 PT used to develop prods. for detection, diagnosis and therapy and for  
 PT modulating activity

XX PS Disclosure; Page 43-46; 54pp; English.

XX CC DNA encoding the human nAChR alpha2 subunit was isolated from a  
 CC human thalamus tissue cDNA library using corresp. rat cDNA. The  
 CC insert of one clone obtd. was ligated with the insert of another  
 CC clone to generate a full-length alpha 2 subunit cDNA. The DNA can be  
 CC used to identify function nAChRs. Cells contg. the DNA can be used  
 CC for screening to identify cpds. which modulate the activity of human  
 CC nAChRs. The human nAChR alpha 2 subunit can be used to product  
 CC antibodies which can be used in immunohistochemistry, diagnosis and  
 CC therapy. The nucleic acids can be used for analysing disease states  
 CC and creating animal models.

XX SQ Sequence 529 AA;

Query Match 18.0%; Score 416; DB 16; Length 529;  
 Best Local Similarity 27.6%; Pred. No. 2.3e-33;  
 Matches 109; Conservative 75; Mismatches 153; Indels 58; Gaps 12;

QY 13 GPGCLPILQLML-MGMLMGITSVPGATATADPKNA-----NKKALDR 53  
 DB 2 GPSCPVILSFKLISLWLLIPAGGEEKRPPTAPGDLSPSPALPGGSHTEDE 61  
 QY 54 LHAGLFTNYDSVQVFPQGTPTNVSLKENVVY-----IDIDELNGKLTTHCWLNLNR 106  
 DB 62 Lfkhlfgrynwarpv-----pntsdvivrfglsiaqlldvdeknqmttnvwlkqgs 116  
 QY 107 DEERWQPSQVDNITQITLKSEVWTPQITLFGNGDEGGLMAETQVTLSH---DGHFRMP 163  
 DB 117 dyklirmpadfgnltslvpsenmipdlvlynadg-favltmktkahlstgtvhwvp 175  
 QY 164 PAVTAYCELMLNMPHDKOSCKLRIGSM---GLKVLPENGTARGESLDHDLVQSPDW 220  
 DB 176 palysscsidvtfpfqgnckmkfgswtydkakidleg---meqtdldkdyesgew 231  
 QY 221 EIVDSRAHFVSODY-----YGYMEYTLTAQRSSMYTAVIYTPASCIIVIALSAFWLP 274  
 DB 222 alvatgynskkydcaaeipdvcyafvtrrplpfyctlnllipclliscslvlfyips 291  
 QY 275 HMGGEKIMINGLLIIVIAFLMFAQLLPVLSNNTPLVVIFYSTSLVSTIVEVLV 334  
 DB 292 dc-geklclclslvllstlvtlllletlpsstslvpllgeyillfemifvltslvltvyl 350  
 QY 335 YLATGKHKR-----RLPEALRKLHGHGCTWLLS 364  
 DB 351 nv-----hnrspsthmphwvrgallgcprwllm 381

RESULT 8

AA09021 ID AAW09021 standard; Protein: 529 AA.

XX AC AAW09021;

XX DT 09-APR-1997 (first entry)

XX DE Neuronal nicotinic acetylcholine receptor alpha-2 subunit.

```

XX  KW  Neuronal nicotinic acetylcholine receptor; nAChR; neurotransmitter;
XX  KW  ligand-gated receptor.
XX  OS  Homo sapiens.
XX  PN  M09641876-A1.
XX  PD  27-DEC-1996.
XX  PF  07-JUN-1996; 96WO-US09775.
XX  PR  07-JUN-1995; 95US-0484722.
XX  PA  (SIBI-) SIBIA NEUROSCIENCES INC.
XX  PI  Elliott KJ, Harpold MM;
XX  DR  WPI: 1997-065463/06.
XX  DR  N-PSDB; AAT48235.
XX  PT  Nucleic acids encoding nicotinic acetylcholine receptor sub-units -
XX  PS  used in screening to determine the effect of drugs on the receptor
XX  PS  Disclosure; Page 49-52; 108pp; English.
XX  CC  The alpha-2 subunit (AAW09021) of the human neuronal nicotinic
XX  CC  acetylcholine receptor (nAChR) can be expressed in transformed
XX  CC  host cells carrying alpha-2 subunit DNA (see also AAT48235). Host
XX  CC  cells, esp. mammalian cells or amphibian oocytes, expressing the
XX  CC  recombinant alpha-2 subunit, opt. in combination with other
XX  CC  recombinant alpha and/or beta subunits (see also AAW09018-20,
XX  CC  AAW09022-27) can be used to examine the function of human AChR and
XX  CC  to identify cpds. that modulate its activity.
XX  SQ  Sequence 529 AA:

Query Match 18.0%; Score 416; DB 18; Length 529;
Best Local Similarity 27.6%; Pred. No. 2.3e-33;
Matches 109; Conservative 75; Mismatches 153; Indels 58; Gaps 12;

QY 13 GPGPLPILQML-HQMLMGITSPGATATADPKNA-----NKKLDR 53
DB 2 gpcspvlfstfkislwllltspaggeaekpprapgdpispsptalpqqgshetedr 61
QY 54 LHAGLFTNYSDDVPVQFGPTNVSLDMVTVY-----IDIDELNGKLTTHCWLNR 106
DB 62 lfkhlfgyrwarpv-----pntsdvlyvrfgisagllldvdeknqmmtnvwlkgews 116
QY 107 DEERWQPSOYDNTQTITLKSSEVWPQITLFNGDEGGLMAETQVTLSH--DGHFRMP 163
DB 117 dylrwnpadfgnltirvsemipdivlynnadge-favtmtkahlfstgtvhwvp 175
QY 164 PAYTATACELMLNMPDRKSKLKIGSW---GLKVLPPENGTRAGSLDHDLDVQSPW 220
DB 176 palyksascsldvffpfdqgnckmkfsgwcydkakldleq---meqvtldkdywesgew 231
QY 221 EYDSRAHFYSODY-----YGYMEYLTFAORRSMTTAVIYPPASCIVILALSAFWLPP 274
DB 232 alynatgtynskkydccaelypdyafvirlrplfthllipcollisceltvvtlyps 291
QY 275 HMGEEKIMINGLLIIVIAFLMFAQLPVLNSNTPLVIFYSTSLYSSTIVEVLV 334
DB 292 dc-geklclsvlslstvlftelipstslvpligeyllftmfvltslvltvfv 350
QY 335 YLATGKKR-----RLPEALRKLLHGLGTWLLIS 364
DB 351 nv-----hhrspstlmpwvrgallgcvprwllm 381

```

```

ID  AC  AAG67161 standard; Protein; 450 AA.
XX  AC  AAG67161;
XX  DT  13-NOV-2001 (first entry)
XX  DE  Amino acid sequence of a human 23821 transporter polypeptide.
XX  KW  Human; transporter; 20685; 579; 17114; 23821; 33894; 32613;
XX  KW  vesicular monoamine transporter; neurotransmitter-symporter;
XX  KW  ABC transporter; sulfate transporter; neurological disorder;
XX  KW  central nervous system disorder; Parkinson's disease; depression;
XX  KW  infectious disease; cell proliferative disorder; cancer; blood disorder;
XX  KW  immune disorder; inflammatory disorder; spleen disorder; lung disorder;
XX  KW  Hodgkin's disease; Niemann-Pick disease; chronic bronchitis; ischemia;
XX  KW  colon disorder; cirrhosis; uterus disorder; endometrium disorder;
XX  KW  endometrial stromal tumour; brain disorder; T-cell disorder; anemia;
XX  KW  Sjogren syndrome; skin disorder; lupus erythematosus; heart disorder;
XX  KW  hematopoietic stem cell; Alzheimer's disease; myocardial infarction;
XX  KW  blood vessel; Kawasaki syndrome; red cell disorder; lymphus disorder;
XX  KW  B-cell disorder; kidney disorder; glomerulonephritis; breast disorder;
XX  KW  testis disorder; thyroid disorder; Graves disease; pancreatitis;
XX  KW  skeletal muscle disorder; tumour; pancreas disorder;
XX  KW  small intestine disorder; celiac sprue.
XX  OS  Homo sapiens.
XX  FT  Key
XX  FT  Location/Qualifiers
XX  FT  8..25
XX  FT  Domain
XX  FT  /note="transmembrane domain"
XX  FT  30..446
XX  FT  /note="neurotransmitter-gated ion channel domain"
XX  FT  236..258
XX  FT  Domain
XX  FT  /note="transmembrane domain"
XX  FT  268..286
XX  FT  Domain
XX  FT  /note="transmembrane domain"
XX  FT  301..320
XX  FT  /note="transmembrane domain"
XX  FT  425..444
XX  FT  /note="transmembrane domain"
XX  PN  W0200164875-A2.
XX  PD  07-SEP-2001.
XX  PF  28-FEB-2001; 2001WO-US06374.
XX  PR  29-FEB-2000; 2000US-0185906.
XX  PA  (MILL-) MILLENNIUM PHARM INC.
XX  PI  Glucksmann MA;
XX  DR  WPI: 2001-550178/61.
XX  DR  N-PSDB; AAH75168.
XX  PT  Novel human transporter polypeptides useful for treating and diagnosing
XX  PT  Parkinson's disease, Hodgkin disease, glomerulonephritis, myocardial
XX  PT  infarction, Grave's disease, Alzheimer's disease, anemia, asthma and
XX  PT  tumours -
XX  PS  Claim 9; Fig 19A-B; 259pp; English.
XX  CC  The present sequence represents a human transporter polypeptide. The
XX  CC  specification describes 20685, 579, 17114, 23821, 33894 or 32613 human
XX  CC  transporter polypeptides. The 20685 transporter is similar to vesicular
XX  CC  monoamine transporters. The 579 transporter is similar to ABC
XX  CC  neurotransmitter-symporters. The 17114 transporter is similar to ABC
XX  CC  transporters. The 32613 transporter and polynucleotides are useful for treating
XX  CC  and diagnosing neurological and central nervous system disorders (e.g.
XX  CC  Parkinson's disease, depression, pain), infectious disease, cell
XX  CC  proliferative disorders (e.g., cancer), blood disorders, and immune and

```

CC inflammatory disorders. They are also useful for treating and  
 CC diagnosing disorders involving the spleen (e.g., Hodgkin disease,  
 CC Niemann-Pick disease), lung (e.g., chronic bronchitis), colon  
 CC (cirrhosis), uterus and endometrium (e.g., endometrial stromal tumours),  
 CC brain (e.g., ischemia), T-cells (e.g., Sjogren syndrome), skin (lupus  
 CC erythematosus), haematopoietic stem cells (e.g., Alzheimer's disease),  
 CC heart (e.g., myocardial infarction), blood vessels (e.g., Kawasaki  
 CC syndrome), red cells (e.g., anemias), disorders involving thymus,  
 CC B-cells, kidney (e.g., glomerulonephritis), disorders involving breast,  
 CC testis, epididymis, prostate, thyroid (e.g., Graves disease), disorders  
 CC involving skeletal muscle (e.g., tumour), pancreas (e.g., pancreatitis),  
 CC small intestine (e.g., celiac sprue), disorders related to reduced  
 CC platelet number and ovary.

XX  
 XX  
 SO Sequence 450 AA:

Query Match 17.8%; Score 412; DB 22; Length 450;  
 Best Local Similarity 27.9%; Pred. No. 4.6e-33;  
 Matches 127; Conservative 78; Mismatches 188; Indels 62; Gaps 18;

QY 24 MGLMLGTSVPGATATADPKANVKALDRHLAGFTNVDSPVYFGCTPT-NVSLKAV 82  
 DB 10 IGIILIFL-IPAECLGAEGRAL-----KIFRDLFANYTSALRPVADTQGLNVLT 62  
 QY 83 VT-YIDIDELNGKLTTHCMLNLRMRDERVWQPSQYDNTOTTLKSEVWFOITLFNGD 141  
 DB 63 Isgitdmderngvltlylwirgwdtdaylrwqpnayggidafrpslswrpdvlylnka 122  
 QY 142 EGGI--MAETQVTLSDHGFRMRPRAVYAYACELMLMWRPHDKQSKLKGSM--GLKV 196  
 DB 123 daqpgsaslnvltlrhdgavtrwdaipatrtscrvaafpfdaqngcltfgwtghqbl 182  
 QY 197 -VLPENGARGLSDHDLVQSPWEIYDSRA-----HFVSQDYGYGMEYTLTQORSS 249  
 DB 183 dvrp-----rgaaasladfvenvewrvtlmparrvtlygcscsepydvtflllrrraa 237  
 QY 250 MYTAVITPASCIVILALSAFWLPRHMGCEKIMINGLITVIAAFMYFAQLPLVLSNNT 309  
 DB 228 ayyvcllllpvcvllslaplafhlpad-sgekvisgtvtialatvqlllaesmp-aeav 295  
 QY 310 PLVVIYSTSLLYSVSTIVEVLVYL-ATGKHKRRPEALRKLLHGLGTWLLSVFST 368  
 DB 296 pllgkyumatmvtfstaltlllmhlycgsrvprpwaralllghnarglcv---re 352  
 QY 369 TGESQAE-KTKEMDHPYFEADEQESSPLG-----INHEVPGAKANQF-- 411  
 DB 353 rgepcgqsrpelpsp-qspegagppapgcpcpcrcicrgaallhv---atlantfrs 408  
 QY 412 -----DNALLATAVDRISEVFSLSAFLIAL 437  
 DB 409 hraagrchedwkrllarvmdrtflalfismaIvmsl 443

RESULT 10  
 AAU00403  
 ID AAU00403 standard; Protein: 450 AA.

XX  
 XX  
 DT 12-SEP-2001 (first entry)

DE Human alpha nicotinic acetylcholine receptor amino acid sequence.

XX Human; alpha nicotinic acetylcholine receptor; alpha10AChR; asthma;  
 KW chronic obstructive lung disease; adult respiratory distress syndrome;  
 KW sepsis; rheumatoid arthritis; osteo-arthritis; Crohn's disease;  
 KW inflammatory bowel disorder; psoriasis; myasthenia gravis; schizophrenia;  
 KW epilepsy; Parkinson's disease; Alzheimer's disease; Tourette's syndrome;  
 KW chronic pain; nicotine addiction; hormonal deficiency.

XX Homo sapiens.

OS  
 XX

PN W0200119973-A2.

XX 22-MAR-2001.

PF 14-SEP-2000; 2000WO-EP09115.

XX 15-SEP-1999; 99US-0153948.

PR 02-FEB-2000; 2000GB-0002431.

XX (JANC) JANSSEN PHARM NV.

XX Yon JR, Grantham CJ, Groot-Kormelink PJ;

DR WPI: 2001-244795/25.

DR N-PSDB: AAS01280.

PT Novel human alpha nicotinic acetylcholine receptor (alpha10AChR) useful  
 PT for identifying modulator compounds for treating psoriasis,  
 PT schizophrenia, asthma, Alzheimer's disease, chronic pain, and nicotine  
 PT addition

XX Claim 4: Fig 6; 72pp: English.

CC The sequence represents the amino acid sequence of human alpha  
 CC nicotinic acetylcholine receptor (Alpha10AChR). The nucleic acid and  
 CC proteins can be used in identifying modulator compounds useful in the  
 CC treatment of asthma, chronic obstructive lung disease, acute adult  
 CC respiratory distress syndrome, sepsis, rheumatoid and osteo-arthritis,  
 CC inflammatory bowel disorder, Crohn's disease and psoriasis, myasthenia  
 CC gravis, schizophrenia, epilepsy, Parkinson's disease, Alzheimer's  
 CC disease, Tourette's syndrome, chronic pain and nicotine addiction.  
 CC The nucleic acid can be used in a test for detecting alpha10AChR encoding  
 CC sequences in the human body or tissues or samples obtained from humans.  
 CC They are also useful in tissue distribution studies in human health and  
 CC disease. Overexpression of the protein in cell lines may be used to  
 CC produce membrane preparations bearing the protein for ligand binding  
 CC studies. The protein is useful as an immunogen to obtain specific  
 CC antibodies which are useful in purification and manipulation of  
 CC polypeptides. The modulator compounds may be used to modulate the  
 CC ion channel activity of alpha10AChR. The modulator compounds also have  
 CC therapeutic value in diseases caused by hormonal deficiency or  
 CC over-production of hormones.

XX Sequence 450 AA:

Query Match 17.7%; Score 409; DB 22; Length 450;  
 Best Local Similarity 27.9%; Pred. No. 9.4e-33;  
 Matches 127; Conservative 78; Mismatches 188; Indels 62; Gaps 18;

QY 24 MGLMLGTSVPGATATADPKANVKALDRHLAGFTNVDSPVYFGCTPT-NVSLKAV 82  
 DB 10 IGIILIFL-IPAECLGAEGRAL-----KIFRDLFANYTSALRPVADTQGLNVLT 62  
 QY 83 VT-YIDIDELNGKLTTHCMLNLRMRDERVWQPSQYDNTOTTLKSEVWFOITLFNGD 141  
 DB 63 Isgitdmderngvltlylwirgwdtdaylrwqpnayggidafrpslswrpdvlylnka 122  
 QY 142 EGGI--MAETQVTLSDHGFRMRPRAVYAYACELMLMWRPHDKQSKLKGSM--GLKV 196  
 DB 123 daqpgsaslnvltlrhdgavtrwdaipatrtscrvaafpfdaqngcltfgwtghqbl 182  
 QY 197 -VLPENGARGLSDHDLVQSPWEIYDSRA-----HFVSQDYGYGMEYTLTQORSS 249  
 DB 183 dvrp-----rgaaasladfvenvewrvtlmparrvtlygcscsepydvtflllrrraa 237  
 QY 250 MYTAVITPASCIVILALSAFWLPRHMGCEKIMINGLITVIAAFMYFAQLPLVLSNNT 309  
 DB 228 ayyvcllllpvcvllslaplafhlpad-sgekvisgtvtialatvqlllaesmp-aeav 295  
 QY 310 PLVVIYSTSLLYSVSTIVEVLVYL-ATGKHKRRPEALRKLLHGLGTWLLSVFST 368  
 DB 296 pllgkyumatmvtfstaltlllmhlycgsrvprpwaralllghnarglcv---re 352

QY	369	TEESQAE-RTKREDDPEYEADEQESSPLG-----INTERVGAKANOF--	411
Db	353	rpepgqgsrppelssps dspeggagppapchepcrlcrgaalnhv---atlantlfrs	408
OY	412	-----DMALLATAVDRISFVSFLAFLIAI	437
Db	409	hraagrchedwkrarvmrdgrfflatfsmalvmsl	443
RESULT	11		
AAE00280			
ID	AAE00280	standard; Protein: 450 AA.	
XX	AAE00280;		
DT	19-JUN-2001	(first entry)	
XX			
DE	Human	nicotinic acetylcholine receptor subunit HNRA10.	
XX			
KW	Human:	nicotinic acetylcholine receptor subunit; HNRA10;	
KW	cytostatic; Alzheimer's disease; memory disorder; anaesthesia; pain;		
KW	epilepsy; schizophrenia; chorea; Tourette's syndrome; depression;		
KW	nicotine dependence; muscular dystrophy; Parkinson's disease; stroke;		
KW	cancer; gastrointestinal ulcer; gastrointestinal motility disorder;		
KW	endocrinal dysfunction; neuronal degenerative disease; gene therapy;		
KW	appetite disorder; flunitus; auditory dysfunction; analgesia.		
KW	nACR; chromosome 12q12;		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Peptide	1..24	
FT		/label= Signal_peptide	
FT	Protein	25..450	
FT		/label= Mature_HNRA10	
FT	Region	/note= "Human nicotinic acetylcholine receptor subunit"	
FT		84..101	
FT		/label= Antigenic_region	
FT	Region	/note= "this region is specifically referred in claim 15"	
FT		120..158	
FT		/label= Antigenic_region	
FT	Region	/note= "this region is specifically referred in claim 15"	
FT		177..196	
FT		/label= Antigenic_region	
FT	Region	/note= "this region and in particular region 179-196 are specifically referred in claim 15"	
FT		207..226	
FT		/label= Antigenic_region	
FT	Domain	/note= "this region is specifically referred in claim 15"	
FT		236..258	
FT		/note= "transmembrane domain/Membrane spanning region (MSR) I"	
FT	Domain	268..285	
FT		/note= "transmembrane domain/Membrane spanning region (MSR) II"	
FT	Domain	298..320	
FT		/note= "transmembrane domain/Membrane spanning region (MSR) III"	
FT	Region	348..395	
FT		/label= Antigenic_region	
FT	Region	/note= "this region and in particular region 357-374 are specifically referred in claim 15"	
FT		405..427	
FT		/label= Antigenic_region	
FT	Domain	/note= "this region is specifically referred in claim 15"	
FT		429..447	
FT		/note= "transmembrane domain/Membrane spanning region (MSR) IV"	
XX			
NN	WO200123551-A1.		
XX			
DD	05-APR-2001.		

xx	PF	08- AUG-2000; 2000MO-EP0718.	
xx	PR	29-SEP-1999; 99EP-0402371.	
xx	PA	(SNPT ) SANOPT-SYNTHELABO.	
xx	PI	Besnard F, Charpentier E, Sgard F;	
xx	DR	WPI: 2001-266152/27.	
xx	DR	N-PSDB: AAD03528.	
xx	XX		
xx	XX	Claim 10; Fig 1; 63pp; English.	
xx	PS		
cc	CC	The present sequence is human nicotinic acetylcholine receptor	
cc	CC	subunit, HNA10. The HNA10 gene is located on human	
cc	CC	chromosome 12q12. HNA10 forms a functional acetylcholine receptor	
cc	CC	only when it is associated with alpha9 subunit. HNA10 is useful	
cc	CC	for identifying agonists and antagonists of alpha9/alpha10 receptor	
cc	CC	(formed by association of alpha9 and HNA10 subunits) which are	
cc	CC	used in pharmaceutical compositions for preventing and treating	
cc	CC	Alzheimer's disease, memory disorders, epilepsy, schizophrenia, pain,	
cc	CC	chorea, Tourette's syndrome, depression, nicotine dependence, cancer,	
cc	CC	muscular dystrophy, Parkinson's disease, gastrointestinal ulcer,	
cc	CC	gastrointestinal motility disorder, endocrinal dysfunction, neuronal	
cc	CC	degenerative disease, stroke, appetite disorder, tinnitus or auditory	
cc	CC	dysfunction or for producing analgesia or anaesthesia. The HNA10	
cc	CC	protein and DNA are useful as research reagents and materials for	
cc	CC	discovery of treatments and diagnostics to animal and human diseases.	
cc	CC	HNA10 DNA is also useful in gene therapy and for chromosome	
cc	CC	identification.	
xx	xx		
SO	Sequence	450 AA:	
	Query Match	17.7%; Score 409; DB 22; Length 450;	
	Best Local Similarity	27.9%; Pred. No. 9.4e-33;	
	Matches 127; Conservative	78; Mismatches 188; Indels 62; Gaps	18;
QY	24	MGMLINGLTSGVGAATADPKANANKALDRHAGLFTNTDSDVQVFGCTPT-NVSLKMY 82	
DB	10	IGIIIIIFL-I-IPaeolgaegrial-----KifrdIfanytsaIrpvadtdqIlnvltE 62	
QY	83	VF-YIDIEPLNKLTTTHGWLNRWDEERWQPSOYDNTOTTLKSSFWPQITLFND 141	
DB	63	ISqIdmderngvLIlyIwIrgewtdaYIrwDnaYggIdaIrpssIswrpdIylnka 122	
QY	142	EGLL-MAETQYTLSDHGHHFRWMPRAVYTYACELNMLMWPMDKQSCKLIGS---GLKV 196	
DB	123	daqppgsastlnvLIrhdgavrtadpaItrscrtvdaaIrfdaqncglIftgswtbgbnj 182	
QY	197	-VLPENGARGLSDHDDLYQSPMEIYDSRA-----HFVSQDYVGKMYETLTNQRSS 249	
DB	163	dVrP-----rgaaasIadIvenvevtrvIgmprarrvlylgccsepydptdIILrrraa 237	
QY	250	MTTAVIYTPASGIVTIALSAFWLPDHMGGEKIMINGLITVIAAFLMFAQLLPVLSNNT 309	
DB	238	ayvcnallIpcvIIsIlaPaflnIpad-egkxslgvtvLIatIrtqIIlaaemp-aesy 295	
QY	310	PLVVIFFSTSLIYSVSTIVEVLYL-ATGKHKRRRLPEALKKLLHGILGTWLLISVST 368	
DB	296	PIlgkyymatmIvftaIatIImIlnlycgsvrIvpawaraIlliIghIargIcv---re 352	
QY	369	TGESQAE-KTRKMDHPREEADEQSSPLG-----INHHVEPAAKNQF-- 411	
DB	353	rgepcgqstppelIspP-qspEgaagppaagchepIrcIcrgaalIinhv---atIentfIs 408	
QY	412	-----DWALLTAVDRISFVSFSLAFLIAT 437	

Db	409	hnaqrchedkkrlyarwmrfflalfafsmalvmsl	443
RESULT	12		
ID	AAP50361		
XX	AAP50361 standard; Protein: 457 AA.		
XX			
AC	AAP50361;		
XX			
DT	08-JAN-1992 (first entry)		
XX			
DE	Human acetyl choline receptor alpha.		
XX			
KW	hACR-alpha; myasthenia gravis; MG; neuromyopathy.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Active-site	42..47	
FT		/label="Epitope	
FT		/note="Claim 1 (I)"	
FT	Active-site	86..91	
FT		/label="Epitope	
FT		/note="Claim 1 (II)"	
FT	Active-site	158..163	
FT		/label="Epitope	
FT		/note="Claim 1 (III)"	
FT	Active-site	181..186	
FT		/label="Epitope	
FT		/note="Claim 1 (IV)"	
FT	Active-site	198..203	
FT		/label="Epitope	
FT		/note="Claim 1 (V)"	
FT	Active-site	354..359	
FT		/label="Epitope	
FT		/note="Claim 1 (VI)"	
FT	Active-site	407..412	
FT		/label="Epitope	
FT		/note="Claim 1 (VII)"	
XX			
PN	JP60078996-A.		
XX			
PD	04-MAY-1985.		
XX			
PF	05-OCT-1983; 83JP-0186402.		
XX			
PR	05-OCT-1983; 83JP-0186402.		
XX			
PA	(MITU ) MITSUBISHI CHEM IND KK.		
XX			
DR	WI; 1985-144120/24.		
XX			
DR	N-PSDB; AAN50415.		
XX			
PT	New peptide(s) useful in treatment of myasthenia gravis - also in diagnosis, having acetyl:choline receptor-alpha like activities.		
XX			
PS	Disclosure; Flg 1; 13pp; Japanese.		
XX			
CC	Peptides derived from the hACR-alpha gene product have similar immunogenicity to the gene product, and may be useful in the treatment and diagnosis of myasthenia gravis caused by neuropathy in autoimmune reaction to ACR.		
CC			
CC			
XX			
Sequence	457 AA;		

[illegible]

CC	AA556386	RESULT 13
XX	AA556386	
XX	AA556386	
AC	AA556386	Standard; Protein; 457 AA.
XX	AA556386	
DT	14-FEB-2000	(first entry)
XX		
DE	Human acetylcholine receptor protein sequence.	
XX		
KW	Human; acetylcholine receptor; AChR; immune response; allergy;	
KW	antibody-mediated disease; gene replacement therapy; T cell epitope;	
KW	dermatological; immunosuppressive; antiinflammatory; haemostatic;	
KW	antianemic; antiallergic; antiaustmatic; antithyroid; antidiabetic;	
KW	autoimmune disease; allergic disease.	
XX		
OS	Homo sapiens.	
XX		
PN	WO930736-A2.	
XX		
PD	24-JUN-1999.	
XX		
PF	16-DEC-1998; 98WO-US26787.	
XX		
PR	16-DEC-1997; 97US-0991143.	
XX		
PA	(MINU ) UNIV MINNESOTA.	
XX		
PI	Contl-Fine BM;	
XX		
DR	WPI: 2000-038343/03.	
DR	N-PSDB; AA38821.	
XX		
PT	Use of T cell epitope peptides for, e.g. preventing allergies -	
XX		
PS	Example 2; Page 219-221; 221pp; English.	
XX		
CC	The present invention describes a method for preventing or inhibiting an	
CC	indication of disease associated with aberrant, pathogenic or undesirable	
CC	antibody production, particularly autoimmune or allergic diseases. The	
CC	method comprises administering a T cell epitope peptide. The sequence of	
CC	the epitope peptide comprises an immunodominant T cell epitope sequence	

CC which is less than the sequence of the antigen. The antigen comprises  
 CC the immunodominant T cell epitope sequence. The method can be used to  
 CC specifically tolerate or down regulate the priming or activity of  
 CC antigen-specific T cells of a mammal. It can be used to prevent or  
 CC inhibit an indication or disease associated with antibody production to  
 CC an antigen such as an endogenous antigen, e.g. acetylcholine receptor,  
 CC insulin, growth hormone, factor VII or factor IX, or an exogenous  
 CC antigen e.g. a fungal antigen, a plant antigen, an antigen of a domestic  
 CC cat or an antigen of a mite. It can be used to treat autoimmune diseases,  
 CC e.g. myasthenia gravis, systemic lupus erythematosus (SLE), Grave's  
 CC disease, autoimmune haemolytic anaemia, autoimmune thrombocytopenia,  
 CC autoimmune asthma, cryoglobulinaemia, thrombotic thrombocytopenic purpura,  
 CC primary biliary sclerosis, pernicious anaemia or pemphigus. It can be  
 CC used for treating allergic diseases, e.g. allergic rhinitis, allergic  
 CC asthma, atopic dermatitis, allergic gastroenteropathy, anaphylaxis,  
 CC urticaria or angioedema. It can also be used in gene therapy for treating  
 CC a disease such as haemophilia or diabetes or an indication such as  
 CC adenosine deaminase deficiency, growth hormone deficiency, insulin  
 CC deficiency, factor IX deficiency or factor VIII deficiency. The present  
 CC sequence represents human acetylcholine receptor which is used in the  
 CC exemplification of the present invention.

XX Sequence 457 AA:

Query Match 17.5%; Score 405.5; DB 21; Length 457;  
 Best Local Similarity 26.9%; Pred. No. 2.2e-32;  
 Matches 116; Conservative 84; Mismatches 165; Indels 67; Gaps 16;

QY 53 RLHAGLFTNDSVOPV---FQCTPTNVSLEMYVYIDIDELNGKLTTHCMLNRDEE 109  
 DB 26 RLVAKLFKDYSSVVRVEDHRYVEVTVGLQL-IQLINDEVNGIVLTNRIKQWADYN 84  
 QY 110 RVMQPSQYDNIITQITLKSSPWTPOITLFGDGC--GLMAEQVTLSDGHFRMPAPVY 167  
 DB 85 LKMPDDYGGVKKIHPSEKIVPDIIVYNAGDFAIVKTVLLIYGYGHICWCPAIF 144  
 QY 168 TAYCELMLNMPHDKOSCKLIGSM---GLKVLPENGTARGESLDHDLVQSPWEIY 223  
 DB 145 KYSCELVTHFPTDEGNSCKLIGTWYDGSVAINPES-----DQPLINFMESGEWV 199  
 QY 224 DSAHNVSDY-----GYMEYTLAQRSSMTAVITPASCIVITLALSFMLPRM 276  
 DB 200 ESTGWKHSVLYSCPTPYDIDYHFWMQPLIYIVNVNIPCLISFGLYGLVPLD- 258  
 QY 277 GGEKINGLILIVIAFLVFAQLPVLNNTPLVIVFSTSLKYSTIVEVLVYL 336  
 DB 259 SGKMTLSISVLSIVLIVLVEIIPSTSAVPLIGKMLFMVFVIAIILTVI-- 316  
 QY 337 ATGKHRR---LPEALKRLHGLGTWLLSVSTTGESQAETKEMDENHY-EEAD 390  
 DB 317 --thhrspsthvmpwvkrvf---ldtlpnlmfstmrkprrek---gdkkltfededl 368  
 QY 391 QESS-----PLG-----INHTVPGA-----KANQ-----FDWALLAVADR 422  
 DB 369 Sdtsqkppppmgfphlpilkhpevksaigikyiaetmksdgsnnaaewkvyamwmdh 428  
 QY 423 ISFVSFSLAFLI 434  
 DB 429 ILLGVFMVLCII 440

RESULT 14  
 AAU00402  
 ID AAU00402 standard: Protein; 433 AA.

XX AAU00402;  
 AC  
 XX 12-SEP-2001 (first entry)  
 DT  
 XX Mature human alpha nicotinic acetylcholine receptor amino acid sequence.  
 DE  
 XX Human; alpha nicotinic acetylcholine receptor; alpha10AChr; asthma;  
 KW

KW chronic obstructive lung disease; adult respiratory distress syndrome;  
 KW sepsis; rheumatoid arthritis; osteo-arthritis; Crohn's disease;  
 KW inflammatory bowel disorder; psoriasis; myasthenia gravis; schizophrenia;  
 KW epilepsy; Parkinson's disease; Alzheimer's disease; Tourette's syndrome;  
 KW chronic pain; nicotine addiction; hormonal deficiency.

OS Homo sapiens.  
 PN WO200119973-A2.

XX 22-MAR-2001.

XX 14-SEP-2000; 2000MO-EP09115.

XX 15-SEP-1999; 99US-0153948.

XX 02-FEB-2000; 2000CB-0002431.

XX (JANC ) JANSSEN PHARM NV.

PI Yon JR, Grantham CJ, Groot-Kormelink PJ;

DR WPI: 2001-244795/25.

DR N-PsDB; AAS01279.

PT Novel human alpha nicotinic acetylcholine receptor (alpha10AChr) useful  
 PT for identifying modulator compounds for treating psoriasis,  
 PT schizophrenia, asthma, Alzheimer's disease, chronic pain, and nicotine  
 PT addition -

XX Claim 4; Page 71-72; 72pp; English.

XX The sequence represents the amino acid sequence of mature human alpha  
 CC nicotinic acetylcholine receptor (Alpha10AChr). The nucleic acid and  
 CC proteins can be used in identifying modulator compounds useful in the  
 CC treatment of asthma, chronic obstructive lung disease, acute adult  
 CC respiratory distress syndrome, sepsis, rheumatoid and osteo-arthritis,  
 CC inflammatory bowel disorder, Crohn's disease and psoriasis, myasthenia  
 CC gravis, schizophrenia, epilepsy, Parkinson's disease, Alzheimer's  
 CC disease, Tourette's syndrome, chronic pain and nicotine addiction.  
 CC The nucleic acid can be used in a test for detecting alpha10AChr encoding  
 CC sequences in the human body or tissues or samples obtained from humans.  
 CC They are also useful in tissue distribution studies in human health and  
 CC disease. Overexpression of the protein in cell lines may be used to  
 CC produce membrane preparations bearing the protein for ligand binding  
 CC studies. The protein is useful as an immunogen to obtain specific  
 CC antibodies which are useful in purification and manipulation of  
 CC polypeptides. The modulator compounds may be used to modulate the  
 CC ion channel activity of alpha10AChr. The modulator compounds also have  
 CC therapeutic value in diseases caused by hormonal deficiency or  
 CC over-production of hormones.

XX Sequence 433 AA:

Query Match 17.5%; Score 404.5; DB 22; Length 433;  
 Best Local Similarity 28.2%; Pred. No. 2.6e-32;  
 Matches 120; Conservative 72; Mismatches 179; Indels 55; Gaps 16;

QY 53 RLHAGLFTNDSVOPVFOGCTPT-NVSLMYVY-IDIDELNGKLTTHCMLNRDEER 110  
 DB 15 klfrdlfanytsairpadtdqclntvleltqslidmdernvltlylwgrewdayl 74  
 QY 111 VMQPSQYDNIITQITLKSSPWTPOITLFGDDEGL--MAEQVTLSDGHFRMPAPVY 168  
 DB 75 rwdpnayggldalrpseslwrpdvlylnkdaqpggsastlnvllrhdgavrvrdapaatr 134  
 QY 169 AYCELMLNMPHDKOSCKLIGSM---GLKV-VLPENGTARGESLDHDLVQSPWEIYD 224  
 DB 135 ssctvdvaafpdaqncglifgswtngbqldvpr-----rgaasladtvenvewrvlg 189  
 QY 225 SRA-----HFVSODYYGYMEYTLAQRSSMTAVITPASCIVITLALSAFLPRHMG 278  
 DB 190 mparrrvlygcsepypdvltflllrrraaayvcnlllpcvllslaplafpads-g 248

**us-09-732-680a-2.ray**

Oy	279	EXIMINGILLIYIAFAIMFEAOLPLPLSNTPLVAFYSLLSYLSTIVELVLYL-A	337
Db	249	ekvsIgvsvellaltvtfglllaeempp-aeavpllgkyymatunvftstcltllimlhy	307
Oy	338	TGKHRRRLPEALRKLLHGLGTWLLSVSTGESQAE-KTKEMDHPYEADQESSPL	396
Db	308	cgpsrvpwpawaralllghlrglcw---rergpcggsrppelpsp-qspeggaqpa	363
Oy	397	G-----INHFEVPGAKNOF-----DVALLATAVDRISFVSLSLA	431
Db	364	gpcheprclcrgaallhvh---atlantlrshraagichedwKrlarvmdrifflaifsm	420
Oy	432	FLILAI 437	
Db	421	alvmsl 426	
RESULT 15			
ID	ABB62694	standard; Protein; 519 AA.	
AC	ABB62694;		
DT	26-MAR-2002	(first entry)	
DE	Drosophila melanogaster	polypeptide SEQ ID NO 14874.	
XX	Drosophila:	developmental biology; cell signalling; insecticide;	
KW	pharmaceutical.		
XX	Drosophila melanogaster.		
PN	MO200171042-A2.		
PD	27-SEP-2001.		
PF	23-MAR-2001;	2001WO-US092331.	
PR	23-MAR-2000;	2000US-1916337P.	
PR	11-JUL-2000;	2000US-0614150.	
XX	(PEKE )	PE CORP NY.	
PI	Venter JC, Adams M, Li PMD,	Myers EW;	
DR	WPI: 2001-656860/75.		
DR	N-PSDB; ABL06797.		
PT	New isolated nucleic acid	detection reagent for detecting 1000 or more	
PT	genes from Drosophila	and for elucidating cell signalling and cell-cell	
PT	interactions -		
PS	Disclosure: SEQ ID NO 14874;	21np + Sequence Listing; English.	
CC	The invention relates to an	isolated nucleic acid detection reagent	
CC	capable of detecting 1000 or	more genes from Drosophila. The invention is	
CC	useful in developmental biology	and in elucidating cell signalling and	
CC	cell-cell interactions in	higher eukaryotes for the development of	
CC	insecticides, therapeutics	and pharmaceutical drugs. The invention	
CC	discloses genomic DNA	sequences (ABL01840-ABL16175) and the encoded	
CC	proteins (ABB57737-ABB72072).		
CC	The sequence data for this	patent did not form part of the printed	
CC	specification, but was	obtained in electronic format directly from	
CC	at ftp.wipo.int/pub/published_pct_sequences.		
SQ	Sequence	519 AA;	

```

0Y 27 LLMGTSVPGKATPAAD---PRNAWVKMLDLRIHALELFNNVSDVOPQKSPRT-NVSLSEW 82
Db 6 llcflvllplansctapisfseapdtk---tlydflsnynrlirpvnmmetcltwaglk 62
0Y 83 VT-YVIDIDELNGKLTTHQWMLRMWDERVWOPSOYONINQITLKSESWPTPOLTFPNCG 141
Db 63 lsgllewnllkqymntcltnwvkgvfdyklrwdpsegygveqlyvpsenhivpdtvlvnmw 122
0Y 142 EGG--LMAETQVTLSDHGFRFMRMPRAYAYACELNMLMNRHDKQSKLKIGSW--GLKV 196
Db 123 dgyevclmtkatllkytgevtweprralykscemmnveyfrydeqicfmkfgsvtlygaqy 182
0Y 137 VLPENGTAKGSSL-----DHDDLQVSEPMELVDSRAHFVSODIY-----GIMEYTLRA 244
Db 183 dlkhldqipgsnllqvvgidltelyfswewdllepra-ckneeyptdllepfsditefkltn 241
0Y 245 QRRSSWNAVITYTPASCIVILALAPMLPRMGEKIMINGLITIVAAFLMFAQLDLPV 304
Db 242 trkclfyvnlilvpcvalctflvtyfvlrpsd-sgekvclstslsvstlvtflillaelp 300
0Y 305 LSNNTPLV-VIEFSTSLIYSVSTIVELVLYATATCKHKRRLEPRLAKRILHGLHGLTWL 362
Db 301 tsiaevllgkylillfmilvslsvtvcvlnihirpsstmmrp-lyrkflflmfpkllmm 359
0Y 363 L-----SVFSTGSEQAERTKEMDENHPEADBOESS-----P 395
Db 360 mrrtglytpdyddspngynyneldivrdsisdfspektdagdydnmgnsvdsdnvlp 419
0Y 366 LGINHTTEVPKAKANP-----DWALLATAVDIRISVPSFLA--FLILAI 437
Db 420 rnlpevlgalravrfiaqhlkadkdnelavedekfvswnldrfllwlfclscvfgltaI 479
0Y 438 RC 439
Db 480 ic 481

```

Search completed: July 18, 2002, 10:16:50  
Job time: 5163 sec

Query Match	17.4%;	Score 402.5;	DB 22;	Length 519;
Best Local Similarity	25.5%;	Pred. No. 5.4e-32;		
Matches 123;	Conservative 92;	Mismatches 192;	Indels 75;	Gaps 16

Fri Jul 19 08:26:19 2002

us-09-732-680a-2.rag

---

